

Série A N° 3768

N° d'ordre

4619

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES
DE L'UNIVERSITÉ DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

Jean René BARRA

PREMIÈRE THÈSE

Contribution à l'étude des lois de probabilité empiriques

DEUXIÈME THÈSE

Propositions données par la Faculté

Soutenues le 21 Octobre 1961 devant la Commission d'examen

MM. PÉRÈS

Président

DUGUÉ

FORTET

}

Examineurs



Série A N° 3768

N° d'ordre

4619

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES
DE L'UNIVERSITÉ DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

Jean René BARRA

PREMIÈRE THÈSE

Contribution à l'étude des lois de probabilité empiriques

DEUXIÈME THÈSE

Propositions données par la Faculté

Soutenues le 21 Octobre 1961 devant la Commission d'examen

MM. PÉRÈS

Président

DUGUÉ

FORTET

} *Examineurs*



FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE PARIS

PROFESSEURS

Doyen J. PERES

G. JULIA	T Analyse supérieure et Algèbre supérieure
DE BROGLIE	T Théories physiques
PRENANT	T Anatomie et histol. comparées
PERES	T Mécanique des fluides et applications
BOURCART	T Géogr. physique et Géol. dynamique
PLANTEFOL	T Botanique
GRASSE	T Evolution des êtres organisés
PREVOST	T Chimie organique
BOULIGAND	T Application de l'analyse à la Géométrie
CHAUDRON	T Chimie appliquée
WYART	T Minéralogie
TEISSIER	T Zoologie
MANGENOT	T Biologie végétale
P. AUGER	T Physique quantique et relativité
MONNIER	T Physiologie générale
PIVETEAU	T Paléontologie
ROCARD	T Physique (E.N.S.)
H. CARTAN	T Mathématiques (E.N.S.)
LAFFITTE	T Chimie Générale
FAVARD	T Géométrie supérieure
COULOMB	T Physique du Globe
Mlle COUSIN	T Biologie animale (S.P.C.N.)
CHRETIEN	T Chimie minérale
P. DRACH	T Zoologie
KÄSTLER	T Physique (E.N.S.)
EPHRUSSI	T Génétique
RIVIERE	T Géologie (SPCN) Orsay
GAUTHERET	T Biol. Végétale (PCB)
LUCAS	T Recherches Physiques
A. THOMAS	T Biol. cellulaire
ARNULF	T Optique appliquée
M. MORAND	T Enseignement de Physique
SOLEILLET	T Physique (P.C.B.)
FORTIER	T Mécanique expérimentale des fluides
DANJON	T Astronomie
PETIT	T Biologie maritime
QUENEY	T Météorologie et dynamique atmosphérique
GALLIEN	T Embryologie
EICHHORN	T Biologie végétale (S.P.C.N.)
DE CUGNAC	T Biologie végétale (S.P.C.N.)
Mlle CAUCHOIS	T Chimie Physique
THELLIER	T Physique du Globe
L'HERITIER	T Génétique
GRIVET	T Radioélectricité
PONCIN	T Mécanique générale
DUBREIL	T Arithmétique et théorie des nombres
QUELET	T Chimie organique
CAGNIARD	T Géophysique appliquée
CHAMPETIER	T Chimie macromolécul.
CUVILLIER	T Micropaléontologie
JUNG	T Pétrographie
TRILLAT	T Microscopie et diffract. électronique
WIEMANN	T Chimie organique et structurale
JACQUINOT	T Spectroscopie et physique céleste
VASSY	T Physique de l'atmosphère
DESTOUCHES	T Théories physiques
PRUVOST	T Géologie
AMIEL	T Chimie générale
HOCART	T Minéralogie
J. P. MATHIEU	T Physique (Orsay)
COUTEAUX	T Biol. animale (PCB)
MAY	T Zoologie (Orsay)
CHOQUET	T Théor. des fonctions et Topologie
FELDMANN	T Biol. végétale marine
GUINIER	T Physique des Solides
JOST	T Physiologie comparée
FORTET	T Calcul des prob. et physique mathématique
SCHWARTZ	T Calc. diff. et Intégr.
CHOUARD	T Physiologie végétale
MALAVARD	T Aviation
BRELOT	T Calcul différentiel
NORMANT	T Synthèse organique
BENARD	T Chimie minérale
BUVAT	T Botanique (E.N.S.)
DUGUE	T Statistiques mathém.
SOULAIRAC	T Psychophysiologie

ULRICH	T Physiologie végétale
MARECHAL	T Optique théorique et appliquée
KIRRMANN	T Théories chimiques
CHADEFAUD	T Botanique
Mlle LE BRETON	T Physiol. générale
SALEM	T Mathématiques générales
LELONG	T Mathématiques générales
DEVILLERS	T Anatomie et Histol. comparées
EHRESMANN	T Topologie algébrique
FRANÇON	T Physique (S. P. C. N.)
GLANGEAUD	T Géogr. physique et Géologie dynamique
GODEMENT	T Mathématiques (MGP)
PISOT	T Techniques mathem. de la Physique
ROCH E.	T Géologie
SCHATZMANN	T Astrophysique
TERMIER	T Paléontologie stratigraphique
ZAMANSKY	T Mathématiques générales
LENNUIER	T Physique (M. G. P.)
RIZET	Génétique
ROUTHIER	T Géologie appliquée
Mme TONNELAT	T Relativité
DIXMIER	T Mathématiques (MPC)
SOUCHAY	T Chimie générale
AIGRAIN	T Electrotechn. générale
BRUSSET	Chimie minérale
M. LEVY	T Physique théorique des hautes énergies
Mme CHAIX	Chimie biologique
Mme HUREL-PY	Biol. vég. (P. C. B.)
PIAUX	Chimie (M. P. C.)
BRUN	T Mécanique exper. des Fluides
LEDERER	T Chimie biologique
Mme DUBREIL	T Mathématiques
Mme LE LONG	T Mathem. (E. N. S.)
BELLAIR	T Géologie (S. P. C. N.)
COTTE	T Physique (Orsay)
DUBOIS J. E.	T Chimie organique
LAMOTTE	Zoologie (E. N. S.)
MICHEL André	T Chimie (S. P. C. N.)
OLMER	T Energétique Gle
ROUAULT	T Physique générale
GAUTHIER	T Mécanique appliquée
BARCHEWITZ	T Chimie Physique
BROSSEL	T Physique atomique
BUSER	Physiologie comparée
CAMUS	Physiol. Vég. (3e Cycle)
CASTAING	Physique générale
CURIEN	Minéralogie et cristallographie
MOYSE	T Physiol. Vég. (Orsay)
PANNETIER	Chimie (S. P. C. N.)
POSSOMPES	T Zoologie
PULLMANN	Chimie théorique
TEILLAC	Physique nucléaire et radioactivité
TONNELAT	Biol. physico-chimique
VILLE	T Econométrie
WILLEMART	Chimie (P. C. B.)
DODE	T Chimie (SPCN) Orsay
FREYMANN	T Physique générale
GUINOCHET	T Biol. Vég. (Orsay)
ROLLET	T Chimie (P. C. B.)
Mlle JOSIEN	T Chimie (P. C. B.)
CHEVALLEY	T Géométrie algébrique et théorie des groupes
ARNOULT	T Electronique
CHAPELLE	T Physique (M. P. C.) (Orsay)
DELANGE	T Mathématiques (Orsay)
DENY	T Mathématiques (MGP)
GERMAIN	T Mécanique théorique des fluides
LUCAS G.	T Géologique
ALLARD	Chimie Physique
BERTHELOT	Physique Nucléaire et Radioactivité
BRICARD	Météorologie et Physique de l'Atmosphère
Mme ALBE-FESSARD	Psychophysiol.
FRIEDEL	Physique des Solides
M. JULIA	Etude des molécules organiques complexes
LENDER	Biol. Animale (S. P. C. N.) (Orsay)
MAGAT	Physico-chimie des Rayonnements
Mlle QUINTIN	T Electrochimie
MONOD	T Chimie du métabolisme
BENOIT	T Physiologie animale
DE POSSEL	T Analyse numérique
CHARLOT	T Chimie analytique
HAISSINSKY	T Radiochimie
LEMEE	T Physiologie B. M. P. V.

GUERIN	T Chimie M. P. C.
CHATELET M	Chimie (S. P. C. N.)
JEAN	Physique Nucléaire et Radioactivité
MATTLER	Physique (S. P. C. N.)
MICHEL (Louis)	Physique théorique des hautes énergies
RUMPF	Chimie (P. C. B.)
BERGERARD	Zoologie (Orsay)
BERTEIN	Electronique
BOUSSIÈRES	Chimie des Radioéléments
Mme COUTURE	Physique (P. C. B.)
ELLENBERGER	Géologie Structurale et Géologie appliquée
FRANC	Biologie animale
SCHNELL	Botanique tropicale
STOLKOWSKI	Physiologie générale
ACHER	Chimie Biologique (Orsay)
BAUDOIN	Biologie animale (P. C. B.)
BEAUMONT	Biologie animale (S. P. C. N.)
BLAMONT	Géophysique Ionosphère
BOURREAU	Biologie Végétale S. P. C. N. (Orsay)
CURIE D.	Physique (M. G. P.)
DURAND-DELGA	Géologie
FREON	Chimie (P. C. B.)
HELLER	Physiologie végétale
TORTRAT	Calcul des Probabilités
CAILLEUX	Géologie (S. P. C. N.)
MAGNAN	Physique (M. P. C.)
DAUDEL	Mécanique Ondulatoire appliquée à la chimie théorique et à la Physique Nucléaire
Mlle FOURCROY	Biologie Végétale (S. P. C. N.)
BARRAUD	Physiologie Cellulaire
NIROTT	Evolution des êtres organisés
Mme FOURES-BRUHAT T	Mécan. Analytique et Mécanique céleste
Mlle VEIL	Physiologie Générale
LWOFF	T Microbiologie
SIESTRUNCK	T Mécanique physique et Expérimentale
LACOMBE	T Métallurgie (ORSAY)
BROCHARD	T Physique P. C. B. (ORSAY)
CABANNES	T Mécanique générale
LESIEUR	T Mathématiques M. P. C. (ORSAY)
MAZET	T Mécanique générale (ORSAY)

Secrétaire général R. POULLAIN

A ma femme



Digitized by the Internet Archive
in 2026

Au moment de présenter ce travail, je voudrais qu'il me soit permis de rendre un très respectueux hommage à la mémoire de Monsieur le Professeur Darmois qui a dirigé avec tant de bienveillance mes débuts en calcul des probabilités.

Je présente mes remerciements les plus sincères à Monsieur le Doyen Pérès qui a accepté de présider le jury de cette thèse. Je prie Monsieur le Professeur Dugué de trouver ici l'expression de ma respectueuse gratitude pour la constante et précieuse attention qu'il m'a accordée tant pour ce travail que pour mon service à la Faculté. Je remercie également Monsieur le Professeur Fortet d'avoir bien voulu faire partie du jury et me donner le sujet d'une seconde thèse.

Je remercie enfin Madame Soury, secrétaire de l'Institut de Statistique de l'Université de Paris, pour l'aide si patiente qu'elle m'a apportée dans toutes les difficultés matérielles liées à ce travail.

INTRODUCTION

L'étude de l'information contenue dans n observations supposées indépendantes d'un même élément aléatoire est aussi ancienne que le Calcul des Probabilités lui-même. Notre but est ici, de définir, à la lumière des résultats concernant les variables aléatoires, l'ensemble des propriétés que l'on peut juger nécessaires pour qu'un groupe d'observations, indépendantes ou non, puisse être considéré comme une estimation correcte d'un élément aléatoire.

Soient donc $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et n éléments aléatoires indépendants sur $\Omega : X_1, \dots, X_n$, de loi commune P_x ; on désigne par Φ_x la loi de probabilité de l'élément certain x de Ω (masse 1 en x) et il s'agit d'abord de préciser en quel sens :

$$P_n = \frac{1}{n} \sum_{i=1}^n \Phi_{x_i}$$

loi de probabilité elle-même aléatoire est une estimation de P_x .

Par loi de probabilité nous entendrons toujours simultanément :

1/ la fonctionnelle linéaire sur l'espace des fonctions numériques boréliennes, c'est-à-dire mesurables par rapport à $\mathcal{A}$. On note $P^*(\varphi)$ cette fonctionnelle, où φ est une fonction borélienne, et P la loi de probabilité en question.

2/ la fonction d'ensemble sur $\mathcal{A}$, et, si elle peut être définie, la fonction de répartition F .

Aux chapitres I et II, en rappelant les résultats classiques pour les variables aléatoires, on met de plus en évidence :

1/ la nature de la convergence de P_n vers P_x . On remarque en particulier que $P_n^*(\varphi)$ converge presque sûrement si et seulement

si $P_x^*(\varphi)$ existe, la limite étant alors $P_x^*(\varphi)$. D'autre part la structure des ensembles de fonctions boréliennes sur lesquels la convergence est uniforme, conduit à remarquer une certaine ressemblance entre $P_n^* - P_x^*$ et une distribution d'ordre 1.

2/ la "loi d'erreur", c'est-à-dire la limite en loi de $\sqrt{n} (P_n - P_x)$ quand $n \rightarrow \infty$. On précise le sens de cette limite en loi en considérant l'élément limite comme aléatoire sur le dual algébrique de l'espace $L_{P_x}^2$ des classes de fonctions boréliennes de carré sommable par rapport à P_x . L'étude de la loi de cet élément aléatoire limite est ramenée à celle d'une fonction aléatoire laplacienne, appelée fonction déterminante d'erreur.

Toujours dans le cas de la droite, on utilise au chapitre III le plan précédemment défini pour étudier le produit $P_n \times Q_m$ (et plus généralement $P_{n_1}^1 \times P_{n_2}^2 \times \dots \times P_{n_k}^k$) comme estimation de $P \times Q$ (et plus généralement de $P^1 \times P^2 \times \dots \times P^k$), $P \times Q$ désignant la loi de probabilité sur R^2 , produit des lois sur R , P et Q . Ceci conduit à établir au chapitre IV des lois des grands nombres pour des variables aléatoires non indépendantes et au chapitre V un théorème Central Limite pour ces mêmes variables aléatoires. On constate que pour ces "observations" non indépendantes $P_n \times Q_m$ possède toutes les propriétés 1/ et 2/ trouvées aux chapitres I et II ; la covariance de la fonction déterminante d'erreur satisfait même à une remarquable propriété minimale.

Ces résultats s'étendent aux vecteurs aléatoires, sauf en ce qui concerne la fonction déterminante d'erreur, dont l'étude conduit à résoudre une équation intégrale. Cette résolution est faite par la méthode de Fredholm et le résultat fait apparaître une relation d'équivalence sur les lois de probabilité sur R^k . Ceci fournit en particulier la généralisation du théorème de Von Mises à des fonctions de répartition empiriques sur R^k . Le rôle important joué dans cette généralisation par la relation d'ordre sur R^k , laisse entrevoir la possibilité de généralisations, par exemple à des lois de probabilité sur un E.V.T. ordonné, mais l'étude des rapports qui devront exister entre la topologie d'un tel espace et le σ -corps de Borel que l'on y choisira, n'entre pas dans la présente étude.

CHAPITRE I

LA LOI DE PROBABILITÉ ALÉATOIRE P_n .

Soient $X_1, \dots, X_n$, n variables aléatoires indépendantes de loi commune P_x et posons :

$$P_n = \frac{1}{n} \sum_{j=1}^{j=n} \Phi_{x_j}$$

où Φ_a désigne la loi de probabilité du nombre certain a . On note P_n^* (resp. P_x^*), F_n (resp. F_x) et $\varphi_n(t)$ (resp. $\varphi_x(t)$), les fonctionnelles linéaires, les fonctions de répartition, et les fonctions caractéristiques de P_n (resp. P_x). P_n est une loi de probabilité aléatoire, qu'on appelle quelquefois loi de probabilité empirique associée à P_x ; elle est en effet déterminée par n "observations indépendantes" de la variable aléatoire X de loi P_x .

A partir des résultats connus, nous nous proposons ici d'étudier la convergence de P_n vers P_x ; on notera qu'il s'agit d'un problème bien propre au Calcul des Probabilités, difficile à traduire par un quelconque mode simple de convergence de mesures, aléatoires ici. En particulier, la donnée de la loi de probabilité (empirique ou non) de la variable aléatoire X entraînant la connaissance de la loi de toutes les variables aléatoires telles que $Y = \varphi(X)$, où φ est borélienne, on ne peut pas restreindre à priori le domaine de définition de P_n^* ou de P_x^* et nous retrouverons cette exigence à propos de tous les théorèmes qui suivent.

A - CONVERGENCE SIMPLE DE P_n^* VERS P_x^*

1 - Convergence presque sûre.

Désignons d'une façon générale par C , l'espace des fonctions numériques boréliennes intégrables par rapport à une loi P :

$\int |\varphi| dP < \infty$; P^* est donc définie sur C_p par $P^*(\varphi) = \int \varphi dP$. On peut maintenant énoncer de la façon suivante, un théorème, en partie classique, de la loi des grands nombres :

Théorème -

"Si $\varphi \in C_{P_X}$. $P_n^*(\varphi) \longrightarrow P_X^*(\varphi)$ p.s.

Si $\varphi \notin C_{P_X}$ $P_n(\varphi)$ n'a pas de limite presque sûre".

$Y = \varphi(X)$ est en effet une variable aléatoire du même champ que X dont l'espérance mathématique est indifféremment donnée par les deux expressions :

$$\int \varphi(x) dP_X = \int y dP_Y.$$

On peut donc se ramener à un échantillon de Y et la première partie du théorème est un énoncé de la loi forte des grands nombres établie par Kolmogoroff ; la deuxième partie découle du lemme suivant :

Lemme - " $\{X_i\}$ étant un échantillon, $\frac{1}{n} \sum_1^n X_i$ ne peut converger en probabilité que vers un nombre certain".

La limite presque sûre de $P_n^*(\varphi)$ étant un nombre certain, la réciproque du théorème de Kolmogoroff [2-2] identifiera en effet ce nombre à $P_X^*(\varphi)$, ce qui prouve le théorème.

2 - Démonstration du lemme.

Soit $\psi(t)$ la seconde fonction caractéristique de la loi commune aux variables aléatoires de l'échantillon ; si :

$$S_n = \frac{1}{n} \sum_1^n X_i$$

converge en probabilité vers une variable aléatoire Z , la loi de S_n tend vers celle de Z et donc :

$$n \psi\left(\frac{t}{n}\right) \longrightarrow \psi_Z(t) \text{ uniformément pour } |t| < T.$$

Mais ceci implique que $\psi(t)$ a une dérivée à droite à l'origine égale à $a + ib$ avec a négatif ; $\psi_Z(t)$ est alors égale à $a|t| + ibt$ d'après la symétrie hermitienne d'une fonction caractéristique. La loi limite de S_n , quand elle existe, est donc une loi de CAUCHY.

Supposons donc maintenant que S_n tende en probabilité vers une variable aléatoire Z , de Cauchy, et mettons en défaut le critère de Slutsky. Si k est un entier supérieur à 1, la variable aléatoire :

$$Y_p = S_{k^{p+1}} - S_{k^p} = \left(\frac{1}{k} - 1\right) \left[\frac{1}{k^p} \sum_1^{k^p} X_i - \frac{1}{k^{p+1} - k^p} \sum_{k^{p+1}}^{k^{p+1}} X_i \right]$$

tend en loi vers la variable aléatoire $2a \left(\frac{1}{k} - 1\right) V$, où V est une variable aléatoire de Cauchy réduite ; ainsi :

Prob. $\{|S_m - S_n| > \varepsilon\}$ ne tend pas vers 0 quand m et n tendent indépendamment vers l'infini.

3 - Convergence presque complètement sûre.

Le lemme du paragraphe précédent, joint au théorème de ERDOS [3] ROBBINS et DUGUE [2-1] de loi des grands nombres en convergence presque complète, fournit le théorème suivant : Nous désignons par C_P^2 l'espace des fonctions boréliennes de carré sommable par rapport à une loi P , on a :

Théorème -

"Si $\varphi \in C_{P_X}^2$ $P_n^*(\varphi) \longrightarrow P_X^*(\varphi)$ p.c.s.

Si $\varphi \notin C_{P_X}^2$ $P_n^*(\varphi)$ n'a pas de limite presque complètement sûre".

4 - Convergence en probabilité.

Appelons moyenne (ou espérance mathématique) faible de la variable aléatoire X le nombre a défini par :

$$\varphi_X'(t) \big|_{t=0} = ia ;$$

Ceci est équivalent à :

$$\begin{cases} \lim_{u \rightarrow +\infty} \int_{-u}^{+u} x dP_X = a \\ x \text{ Prob. } \{|X| > x\} \longrightarrow 0 \text{ quand } x \longrightarrow \infty. \end{cases}$$

Il faut noter que cette définition est cohérente pour toutes les variables d'un même champ : soit g une fonction borélienne, et Y la variable aléatoire $g(X)$ de loi P_Y . Si P_Y a une espérance mathématique faible a , on a :

$$x \text{ Prob. } \{|Y| > x\} \longrightarrow 0 \text{ quand } x \longrightarrow +\infty$$

ce que l'on peut écrire :

$$x \text{ Prob. } \{|g(X)| > x\} \longrightarrow 0 \text{ quand } x \longrightarrow +\infty ; \quad (1)$$

d'autre part :

$$\lim_{u \rightarrow +\infty} \int_{-u}^{+u} y \, dP_Y = a$$

ce que l'on écrira encore :

$$\lim_{u \rightarrow +\infty} \int_{|g(x)| < u} g(x) \, dP_X = a \quad (2)$$

Notons $\tilde{E}(X)$ l'espérance mathématique faible de la variable aléatoire X ; (1) et (2) justifient alors que l'on écrive ici, $\tilde{E}\{g(X)\} = a$. Remarquons que si X a une espérance mathématique, il a aussi une espérance mathématique faible et les deux nombres sont égaux. Désignons maintenant par C_P^0 l'espace des fonctions g , boréliennes faiblement intégrables par rapport à une loi P ; $\tilde{P}^*(g)$ est définie sur C_P^0 par $\tilde{P}^*(g) = \tilde{E}\{g(X)\}$ si X a pour loi P . On a par le même procédé qu'aux paragraphes précédents :

Théorème -

$$\begin{aligned} \text{"Si } \varphi \in C_{P_X}^0 \quad P_n^*(\varphi) &\longrightarrow \tilde{P}_n^*(\varphi) \text{ en probabilité,} \\ \text{Si } \varphi \notin C_{P_X}^0 \quad P_n^*(\varphi) &\text{ n'a pas de limite en probabilité".} \end{aligned}$$

On remarque, ici comme d'ailleurs pour les deux théorèmes précédents, que l'espace aléatoire C_{P_n} contient, presque sûrement quel que soit n , l'espace des fonctions boréliennes bornées sur tout compact du support de P .

5 - Propriétés de l'espérance mathématique faible.

Si $\tilde{E}(X)$ et $\tilde{E}(Y)$ existent, on a :

$$\tilde{E}(X + Y) = \tilde{E}(X) + \tilde{E}(Y) ;$$

ceci résulte en effet directement du théorème précédent :

Si $\frac{X_1 + \dots + X_n}{n} \rightarrow a$ et $\frac{Y_1 + \dots + Y_n}{n} \rightarrow b$, en probabilité, quelle que soit la corrélation entre X et Y ,

$$\frac{(X_1 + Y_1) + (X_2 + Y_2) + \dots + (X_n + Y_n)}{n} \rightarrow a + b$$

en probabilité. On peut aussi retrouver ce résultat en considérant la loi du couple (X, Y) et en appliquant les critères du § 4.

Par contre en général XY n'a pas une espérance mathématique faible égale à $\tilde{E}(X) \cdot \tilde{E}(Y)$ si X et Y sont indépendantes et ont seulement une espérance mathématique faible, comme le prouve le contre-exemple suivant :

Proposition - "Soient X et Y deux variables aléatoires indépendantes de même loi $F(x)$, symétrique et définie par :

$$F(x) = 1 - \frac{1}{x(\log x)^\alpha} \quad \text{pour } x > a > 1 \quad \left(0 < \alpha < \frac{1}{2}\right)$$

la variable aléatoire $Z = XY$ a une fonction de répartition $G(z)$ telle que :

$$z \text{ Prob. } \{|Z| > z\} \rightarrow +\infty \quad \text{si } z \rightarrow +\infty.$$

Posons pour simplifier l'écriture $f(x) = (\log x)^\alpha$; si $z > a^2$ on a :

$$\text{Prob. } \{|Z| > z\} = 4 \int_{\substack{x, y > a \\ xy > z}} d\left(\frac{1}{xf(x)}\right) d\left(\frac{1}{yf(y)}\right) = 4 P(z)$$

soit encore :

$$P(z) = \int_a^{\frac{z}{a}} \left[\frac{1}{xf(x)} \right]_{\frac{z}{y}}^{\infty} d\left(\frac{1}{yf(y)}\right) + \left[\frac{1}{xf(x)} \right]_a^{\infty} \left[\frac{1}{yf(y)} \right]_{\frac{z}{a}}^{\infty}$$

écrivons :

$$d\left(\frac{1}{yf(y)}\right) = \frac{1}{y} d\left(\frac{1}{f(y)}\right) - \frac{dy}{y^2 f(y)}$$

il vient finalement :

$$z P(z) = - \int_a^{\frac{z}{a}} \frac{1}{f\left(\frac{z}{y}\right)} d\left(\frac{1}{f(y)}\right) + \int_a^{\frac{z}{a}} \frac{dy}{yf(y)f\left(\frac{z}{y}\right)} + \frac{1}{f(a)f\left(\frac{z}{a}\right)}$$

la première intégrale étant bornée par $\frac{1}{f^2(a)}$, on voit aisément que la proposition sera établie si :

$$J = \int_a^z \frac{dy}{y f(y) f\left(\frac{z}{y}\right)} \text{ tend vers } +\infty, \text{ si } z \rightarrow +\infty.$$

Revenons alors à la forme de f , posons $\log a = t_0$, $\log z = t$, il vient :

$$J = \int_{t_0}^{t-t_0} \frac{du}{(t-u)^\alpha u^\alpha} = t^{1-2\alpha} \int_{\frac{t_0}{t}}^{1-\frac{t_0}{t}} \frac{dx}{(1-x)^\alpha x^\alpha}$$

et comme $0 < \alpha < \frac{1}{2}$ $\int_{\frac{t_0}{t}}^{1-\frac{t_0}{t}} \frac{dx}{(1-x)^\alpha x^\alpha}$ tend vers $\int_0^1 \frac{dx}{(1-x)^\alpha x^\alpha}$ qui n'est pas nul, quand $z \rightarrow +\infty$; donc $J \rightarrow +\infty$ comme $t^{1-2\alpha}$ quand $z \rightarrow +\infty$.

On voit sur cette dernière forme qu'il est facile d'obtenir d'autres contre-exemples du même type.

B - CONVERGENCE UNIFORME DE P_n^* VERS P_x^*

1 - On situe le problème en remarquant d'abord que P_n^* ne tend pas vers P_x^* uniformément sur C_{P_x} , au sens d'une convergence du Calcul des Probabilités, en effet :

Théorème - "La variation totale de $F_n(x) - F_x(x)$ tend presque sûrement vers un nombre positif quand $n \rightarrow +\infty$, s'il existe pour F_x un intervalle J de continuité avec $P_x(J) \neq 0$ ".

Sur J , $F_x(x)$ est continue croissante et $F_n(x)$ étagée ; la variation totale de la différence est égale sur J à :

$$\frac{g}{n} + P_x(J)$$

où $\frac{g}{n}$ est la somme des sauts de $F_n(x)$ dans J . Or $\frac{g}{n}$ tend vers $P_x(J)$ quand $n \rightarrow +\infty$.

En particulier on déduit de ceci que, presque sûrement, la convergence de P_n^* vers P_x^* n'est pas uniforme sur l'espace des fonctions caractéristiques d'ensembles boréliens.

2 - On cherchera donc les sous-ensembles de C_{P_X} , sur lesquels on a convergence uniforme ; il y a essentiellement deux théorèmes à la base de cette recherche :

Théorème de GLIVENKO-CANTELLI [2-2].

$$\sup_x |F_n(x) - F_X(x)| \rightarrow 0 \quad \text{p.s. et même p.c.s.}"$$

Théorème de FORTET et MOURIER.

"Si $\alpha(x)$ est une fonction non décroissante de C_{P_X} et r un nombre réel, $r > 1$, on a :

$$\int_{-\infty}^{+\infty} |F_n(x) - F_X(x)|^r d\alpha(x) \rightarrow 0 \quad \text{p.s. quand } n \rightarrow +\infty."$$

On peut se ramener au cas $\alpha(x) = x$, en passant à la nouvelle variable $Y = \alpha(X)$; sous cette forme, ce théorème a été obtenu pour $r = 1$, dans FORTET et MOURIER [4], comme cas particulier de théorèmes concernant des éléments aléatoires dans un espace de Banach. Nous donnons ici une démonstration directe qui établit de plus que si $\alpha(x)$ appartient à $C_{P_X}^2$ la convergence est presque complètement sûre.

Démonstration.

Soit M un nombre positif quelconque pour l'instant ; décomposons l'intégrale :

$$\int_{-\infty}^{+\infty} |F_n(x) - F_X(x)|^r dx,$$

en trois intégrales :

$$J_1 = \int_{-\infty}^{-M} |F_n(x) - F_X(x)|^r dx$$

$$J_2 = \int_{+M}^{+\infty} |F_n(x) - F_X(x)|^r dx$$

$$J_3 = \int_{-M}^{+M} |F_n(x) - F_X(x)|^r dx$$

Pour J_1 on remarque que :

$$|F_n(x) - F_X(x)|^r < 2^{r-1} [(F_n(x))^r + (F_X(x))^r] < 2^{r-1} [F_n(x) + F_X(x)] ;$$

de même on majore J_2 par :

$$2^{r-1} \int_{+M}^{\infty} [2 - F_n(x) - F_x(x)] dx.$$

Posons :

$$Q_n(x) = 1 - F_n(x) + F_n(-x)$$

et :

$$Q_x(x) = 1 - F_x(x) + F_x(-x),$$

Sachant que $x Q_n(x)$ tend vers 0, quand $x \rightarrow +\infty$, puisque $Q_n(x) = 0$ pour x assez grand, et que d'autre part $x Q_x(x)$ tend vers 0 quand $x \rightarrow +\infty$, car $E(X)$ existe, puisque α appartient à C_{p_x} , on intègre par parties dans J_1 et J_2 et il vient :

$$J_1 + J_2 < 2^{r-1} \int_{|x| > M} |x| d(F_n(x) + F_x(x)) + 2^{r-1} \cdot M \cdot [Q_n(M) + Q_x(M)].$$

Quelque soit $\varepsilon > 0$, donné on peut choisir M tel que :

$$M Q_x(M) < \frac{\varepsilon}{2^{r+2}} \quad \text{et que} \quad \int_{|x| > M} |x| dP_x < \frac{\varepsilon}{2^{r+2}},$$

puisque $E(X)$ existe ; M étant maintenant choisi on sait que, d'après le théorème de Kolmogoroff :

$$M Q_n(M) \rightarrow M Q_x(M) \quad \text{p.s.} \quad \text{et} \quad \int_{|x| > M} x dP_n \rightarrow \int_{|x| > M} x dP_x \quad \text{p.s.}$$

quand $n \rightarrow +\infty$,

et que d'autre part, d'après le théorème de Glivenko-Cantelli :

$$J_3 = \int_{-M}^{+M} |F_n(x) - F_x(x)|^r dx \rightarrow 0 \quad \text{p.s.}$$

Ainsi, quel que soit ε , il existe M , donc presque sûrement un $N(\varepsilon)$ tel que :

$$\int_{-\infty}^{+\infty} |F_n(x) - F(x)|^r dx < \varepsilon \quad \text{pour} \quad n > N(\varepsilon)$$

c'est la convergence presque sûre cherchée.

De même la convergence presque complètement sûre résulte de ce qui suit : choisissons ε et M de la même façon que précédemment et écrivons :

$$J_1 + J_2 < 2^{r-1} \int_{|x| > M} |x| d(F_n(x) - F_x(x)) + 2^{r-1} M [Q_n(M) - Q_x(M)] + \frac{\varepsilon}{2}$$

les variables aléatoires :

$$\int_{|x| > M} |x| d(F_n(x) - F_x(x)) \quad \text{et} \quad [M Q_n(M) - Q_x(M)]$$

tendent presque complètement sûrement vers 0, quand $n \rightarrow +\infty$, la première en vertu du théorème de Glivenko-Cantelli, la seconde en vertu de la loi des grands nombres, $E(X^2)$ étant ici supposé fini.

3 - Les seuls théorèmes de convergence uniforme que l'on ait, découlent des deux précédents en considérant les fonctions boréliennes de la forme :

$$\varphi(x) = \int_{x_0}^x g(u) d\alpha(u),$$

où $\alpha(u)$ est une fonction non décroissante et où $g(u)$ est intégrable par rapport à α dans tout compact ; pour une fonction α fixée, on désignera par D_α cet espace fonctionnel.

Théorème I -

" $P_n^*(\varphi)$ tend vers $P_x^*(\varphi)$ quand $n \rightarrow +\infty$ presque sûrement uniformément sur les sous-ensembles de D_α définis par :

$$\int_{-\infty}^{+\infty} |g(u)| d\alpha(u) < M."$$

On suppose donc ici g intégrable par rapport à α sur toute la droite et l'on peut prendre :

$$\varphi(x) = \int_{-\infty}^x g(u) d\alpha(u),$$

d'où :

$$P_n^*(\varphi) - P_x^*(\varphi) = \iint_{u < x} g(u) d\alpha(u) d[F_n(x) - F_x(x)]$$

soit encore :

$$\int_{-\infty}^{+\infty} g(u) [F_n(u) - F_x(u)] d\alpha(u)$$

que l'on majore par :

$$\sup_x |F_n(x) - F_x(x)| \cdot \int_{-\infty}^{+\infty} |g(u)| d\alpha(u),$$

et la proposition découle du théorème de Glivenko-Cantelli ; de même le théorème de Fortet et Mourier donne les deux propositions suivantes :

Théorème II -

"Si α appartient à C_{P_x} , $P_n^*(\varphi)$ tend vers $P_x^*(\varphi)$, quand $n \rightarrow +\infty$, presque sûrement uniformément sur les sous-ensembles de D_α définis par :

$$\sup_u |g(u)| < M."$$

On suppose ici que les fonctions φ sont absolument continues par rapport à α , et on écrit encore :

$$P_n^*(\varphi) - P_x^*(\varphi) = \iint_{u \in (x, x_0)} g(u) d\alpha(u) d[F_n(x) - F_x(x)]$$

soit :

$$\int_{-\infty}^{x_0} g(u) [F_n(u) - F_x(u)] d\alpha(u) - \int_{x_0}^{+\infty} g(u) [F_n(u) - F_x(u)] d\alpha(u)$$

que l'on majore par :

$$M \int_{-\infty}^{+\infty} |F_n(u) - F_x(u)| d\alpha(u),$$

et cette expression tend presque sûrement vers 0, quand $n \rightarrow \infty$.

Théorème III -

"Si α appartient à C_{P_x} , $P_n^*(\varphi)$ tend vers $P_x^*(\varphi)$, quand $n \rightarrow \infty$, presque sûrement uniformément sur les sous-ensembles de D_α définis par :

$$\int_{-\infty}^{+\infty} |g(u)|^r d\alpha(u), \quad (r > 1)".$$

$P_n^*(\varphi) - P_x^*(\varphi)$ est encore ici majorée par :

$$\int_{-\infty}^{+\infty} |g(u)| \cdot |F_n(u) - F_x(u)| d\alpha(u),$$

expression qui est elle-même inférieure à :

$$\left[\int_{-\infty}^{+\infty} |g(u)|^r d\alpha(u) \right]^{\frac{1}{r}} \cdot \left[\int_{-\infty}^{+\infty} |F_n(u) - F_x(u)|^s d\alpha(u) \right]^{\frac{1}{s}}$$

si l'on a : $\frac{1}{r} + \frac{1}{s} = 1$; et nous concluons, ici encore, grâce au théorème de Fortet et Mourier.

Concluons : comme mesure, P_n ne tend presque sûrement pas vers P_x , mais P_n tend presque sûrement vers P_x , comme ce que l'on pourrait appeler une distribution d'ordre 1, par rapport à α ; en effet, si dans les énoncés précédents on fait :

$$\alpha(x) = \begin{cases} K & \text{pour } x > K \\ x & \text{pour } -K \leq x \leq K \\ -K & \text{pour } x < -K, \end{cases}$$

le théorème II est l'énoncé de la convergence presque sûre de P_n vers P_x comme distribution d'ordre 1. Cette remarque dominera encore le chapitre suivant et doit être considérée comme le caractère essentiel de la convergence de P_n vers P_x .

CHAPITRE II

LA LOI D'ERREUR STATISTIQUE Δ_n .

Nous avons vu au chapitre précédent comment P_n tendait vers P_x , quand $n \rightarrow \infty$, nous complétons ici l'étude de P_n en tant qu'estimation de P_x par l'étude de la différence :

$$\Delta_n = \sqrt{n} (P_n - P_x) ;$$

Δ_n n'est pas une loi de probabilité, mais la fonctionnelle linéaire :

$$\Delta_n^* = \sqrt{n} (P_n^* - P_x^*)$$

est définie pour certaines fonctions numériques boréliennes, par exemple, pour tout n , les fonctions boréliennes bornées sur tout compact du support de P_x ; mais C_{P_x} et $C_{P_x}^2$ seront aussi envisagées comme domaines de définition pour Δ_n^* .

Une loi d'erreur statistique sera donc, pour nous ici, la loi de la fonctionnelle linéaire aléatoire, limite, quand elle existe, de la loi de la fonctionnelle Δ_n^* ; c'est en effet une conséquence immédiate du théorème de Slutsky que, lorsque $n \rightarrow +\infty$, et ceci quel que soit φ , $\Delta_n^*(\varphi)$ n'a pas de limite en probabilité ; on se propose donc seulement de définir ici, ce qu'il suffit d'entendre par "loi limite de la fonctionnelle Δ_n^* " et ses propriétés.

A - LIMITE SIMPLE EN LOI DE Δ_n^* DANS $C_{P_x}^2$

1 - $\Delta_n^*(\varphi)$ n'ayant de limite en loi que si φ appartient à $C_{P_x}^2$ (théorème de Laplace), on restreint d'abord le domaine de définition de Δ_n^* à $C_{P_x}^2$.

Soit $L_{P_x}^2$, l'espace de Hilbert des classes d'équivalences de $C_{P_x}^2$ définies par :

$$\varphi_1 \sim \varphi_2 \text{ si } P_x^*(\varphi_1 - \varphi_2)^2 = 0;$$

$L_{P_x}^2$ a une base finie ou dénombrable suivant la nature du support de P . [5].

Pour toutes les fonctions φ d'une même classe de $L_{P_x}^2$, $\Delta_n^*(\varphi)$ a la même loi, mais Δ_n^* n'en est pas pour autant un élément aléatoire de $L_{P_x}^2$, dual algébrique de $L_{P_x}^2$. Nous allons cependant montrer que la limite en "loi temporelle" de Δ_n^* peut être considérée comme la loi d'un élément aléatoire de $L_{P_x}^2$. On sait que, quels que soient $\varphi_1, \varphi_2, \dots, \varphi_r$ appartenant à $L_{P_x}^2$, le vecteur aléatoire $\{\Delta_n^*(\varphi_1), \dots, \Delta_n^*(\varphi_r)\}$ tend en loi vers un vecteur aléatoire normal défini par une certaine matrice de covariance ; on peut exprimer ceci d'une façon différente :

2 - Le projecteur aléatoire de Wiener-Lévy.

Soit l'application linéaire aléatoire de $L_{P_x}^2$ dans R , définie par :

$$\Delta^*(\varphi) = \sum_{k=1}^{\infty} U_k \langle f_k, \varphi \rangle \quad (1)$$

où les U_k sont des variables aléatoires normales indépendantes centrées réduites et où $f_0 = 1, f_1, \dots, f_k, \dots$ est une base de $L_{P_x}^2$.

Pour tout φ , $\Delta^*(\varphi)$ est une variable aléatoire normale d'écart moyen quadratique $\|\varphi\|$, mais Δ^* ne définit pas pour autant une fonctionnelle linéaire aléatoire sur $L_{P_x}^2$, car la série $\sum_k U_k f_k$ est presque sûrement divergente. La loi de cette application linéaire ne dépend pas de la base choisie et pourrait être entièrement donnée par le caractère laplacien joint à l'une des deux conditions :

$$\Delta^* \text{ est linéaire et } E\{\Delta^*(\varphi)^2\} = \|\varphi\|^2$$

ou :

$$\text{la "covariance" : } E\{\Delta^*(\varphi) \Delta^*(\varphi')\} = \langle \varphi, \varphi' \rangle.$$

le résultat précédent peut alors s'énoncer :

Proposition - " Δ_n^* tend en loi temporelle vers Δ^* quand $n \rightarrow \infty$."

On entend ici, par loi temporelle, par analogie avec les fonctions aléatoires, la loi du vecteur :

$\{\Delta^*(\varphi_1), \dots, \Delta^*(\varphi_k)\}$ pour tout k et toute suite $\varphi_1, \dots, \varphi_k$ d'éléments de $L_{P_x}^2$.

Donnons de cette proposition une démonstration directe : Etant donnée la variable aléatoire X , la fonctionnelle linéaire aléatoire Φ_X admet la représentation en loi :

$$\Phi_X(\varphi) = \sum_{k=0}^{\infty} f_k(X) \langle f_k, \varphi \rangle$$

et donc :

$$\Delta_n^*(\varphi) = \sum_{k=1}^{\infty} A_n^k \langle f_k, \varphi \rangle$$

où le système des variables A_n^k est défini par :

$$A_n^k = \frac{1}{\sqrt{n}} \sum_{i=1}^n f_k(X_i)$$

Comme d'autre part on a :

$$E\{f_k(X_i)\} = \langle f_k, f_0 \rangle = 0$$

$$E\{f_k(X_i) f_{k'}(X_i)\} = \langle f_k, f_{k'} \rangle = 0$$

$$\text{et :} \quad E\{f_k(X_i)^2\} = \|f_k\|^2 = 1,$$

la loi du système des A_n^k , tend, quand $n \rightarrow \infty$ vers celle d'un système de variables aléatoires normales centrées réduites et indépendantes.

3 - Δ^* , limite en loi temporelle de Δ_n^* n'est ainsi dans $L_{P_X}^2$ qu'une application linéaire aléatoire, qui ne peut être continue, car cela reviendrait à utiliser la convergence uniforme de P_n^* vers P_X^* ; on doit donc, pour espérer définir une fonctionnelle linéaire aléatoire continue en un certain sens, se restreindre aux fonctions de D_α qui appartiennent aussi à $C_{P_X}^2$, et ceci, comme au chapitre précédent, peut se faire de plusieurs façons.

B - Δ^* FONCTIONNELLE LINEAIRE SUR D_α^1

1 - En accord avec la terminologie de la page 11 désignons par :

l_α^1 l'espace des fonctions g , sommables par rapport à la fonction α , non décroissante sur R :

$$\int_{-\infty}^{+\infty} |g(u)| d\alpha(u) < \infty.$$

et par :

D_α^1 l'espace des fonctions φ , s'écrivant :

$$\varphi(x) = \int_{-\infty}^x g(u) d\alpha(u).$$

où g appartient à L_α^1 .

Les fonctions φ de D_α^1 , sont bornées et appartiennent donc à $C_{P_X}^2$ et l'on a :

$$\Delta^*(\varphi) = \sum_{k=1}^{\infty} U_k \int_{-\infty}^{+\infty} g(u) F_k(u) d\alpha(u)$$

où $F_k(u) = \int_u^{\infty} f_k(x) dP_X$ est bornée quel que soit u .

2 - La série :

$$\sum_{k=1}^{\infty} U_k \int_u^{\infty} f_k(x) dP_X \quad (1)$$

est, quel que soit u , presque sûrement convergente ; il est en effet élémentaire d'établir que quels que soient u et v :

$$\sum_{k=1}^{\infty} \int_u^{\infty} f_k(x) dP_X \int_v^{\infty} f_k(x) dP_X = F_X[\inf(u, v)] - F_X(u) F_X(v). \quad (2)$$

en effet si $\chi_u(x)$ est la fonction caractéristique de l'intervalle (u, ∞) , on a dans $L_{P_X}^2$:

$$\int_u^{\infty} f_k(x) dP_X = \langle \chi_u, f_k \rangle$$

et donc :

$$\sum_{k=1}^{\infty} \langle \chi_u, f_k \rangle \cdot \langle \chi_v, f_k \rangle = \langle \chi_u, \chi_v \rangle - \langle \chi_u, f_0 \rangle \langle \chi_v, f_0 \rangle$$

Soit maintenant $\beta(u)$, une fonction numérique non décroissante telle que :

$$\beta(-\infty) = 0$$

et :

$$\beta^2 \cdot (1 - F_x) \in C_{P_x}$$

et soit L_{β}^2 , l'espace des classes de fonctions de carré sommable par rapport à β ; on a :

Proposition - "La série (1) définit presque sûrement un élément aléatoire laplacien de L_{β}^2 ; cet élément aléatoire a une loi indépendante de la base $\{f_k\}$ de $L_{P_x}^2$, choisie, et cette loi est entièrement déterminée par sa covariance :

$$F_x [\inf(t, \tau)] - F_x(t) F_x(\tau)."$$

L'indépendance du problème, quant à la base $\{f_k\}$ découle de la même propriété pour le projecteur de Wiener-Lévy. Considérons ensuite la série :

$$\sum_{k=0}^{\infty} U_k \int_u^{\infty} f_k(x) dP_x \quad (3)$$

qui ne diffère de (1) que par le premier terme. Il est facile de voir que $\beta[\inf(x, y)]$ est un noyau symétrique défini positif de $L_{P_x}^2$; soient λ_k et $\beta_k(x)$ ses éléments propres, prenons ces β_k comme base particulière $\{f_k\}$ dans (3), bien que β_0 ne soit pas égal à 1; les fonctions :

$$g_k(u) = \frac{1}{\sqrt{\lambda_k}} \int_u^{\infty} \beta_k(x) dP_x$$

forment alors un système orthonormal de L_{β}^2 , d'ailleurs associé au noyau $1 - F_x[\sup(u, v)]$ et la série (3) s'écrit alors :

$$\sum_{k=0}^{\infty} \frac{1}{\sqrt{\lambda_k}} U_k g_k(u)$$

définissant presque sûrement un élément aléatoire de L_{β}^2 puisque la série $\sum_k \frac{1}{\lambda_k} U_k^2$ est presque sûrement convergente $\left(\sum_0^{\infty} \frac{1}{\lambda_k} = P_x^*(\beta) \right)$.

Un changement de base n'affecte pas la loi des U_k , et ainsi (3) définit, quel que soit la base $\{f_k\}$, (f_0 étant ou non égal à 1) une fonction aléatoire $W(u)$ de L_{β}^2 , et il en est alors de même de la série (1). Notons $\Delta(u)$ la somme de la série (1); dans une représentation avec $f_0 = 1$, $W(u)$ est égal à :

$$\Delta(u) + [1 - F(u)] U_0$$

et l'on vérifie que la covariance Γ_W de W :

$$\Gamma_W = 1 - F_x [\sup(u, v)]$$

et la covariance Γ_Δ de Δ :

$$\Gamma_\Delta = F_x [\inf(x, y)] - F_x(x) F_x(y)$$

satisfont bien à :

$$\Gamma_W = \Gamma_\Delta + [1 - F_x(x)] [1 - F_x(y)].$$

3 - On peut remarquer que le procédé précédent ne donne pas pour Δ une représentation sur un système orthonormal de L^2_β , alors qu'il en fournit une pour $W(u)$, ceci parce que β_0 est différent de 1 ; enfin dans la suite on prendra le plus souvent $\beta = F_x$, et on énonce la proposition qui sera démontrée plus loin :

Proposition I - "Il existe une fonction aléatoire $\Delta(u)$, presque sûrement continue et bornée qui est représentée dans $L^2_{F_x}$ par la série (1) ; $\sup_u |\Delta(u)|$ est alors une variable aléatoire suivant la loi de Kolmogoroff."

Ainsi pour φ appartenant à D^1_α , $\Delta^*(\varphi)$ peut s'écrire $\langle g, \Delta \rangle$, où g appartient à l^1_α et $\Delta(u)$ bornée ; on a donc ainsi une véritable fonctionnelle linéaire sur l^1_α .

C - Δ^* FONCTIONNELLE LINEAIRE SUR D^∞_α

1 - Toujours en accord avec la terminologie de la page 11, soit $\alpha(x)$ une fonction numérique non décroissante et appartenant à $C^2_{F_x}$; désignons par :

l^∞_α l'espace des fonctions g bornées et intégrables par rapport à α sur tout compact,

et par :

D^∞_α l'espace des fonctions φ s'écrivant :

$$\int_a^x g(u) d\alpha(u)$$

où $g(u)$ appartient à l^∞_α .

On peut noter que ceci revient à dire que φ est absolument continue par rapport à α et par ailleurs appartient à $C_{p_X}^2$.

2 - On écrit, comme au paragraphe précédent :

$$\Delta^*(\varphi) = \int_a^\infty g(u) \Delta(u) d\alpha(u) + \int_{-\infty}^a g(u) \Delta(-u) d\alpha(u)$$

et on prévoit la proposition :

Proposition II - "Il existe une fonction aléatoire $\Delta(u)$ presque sûrement intégrable par rapport à $\alpha(u)$, qui est représentée dans $L_{p_X}^2$ par la série (1) ; $\int_{-\infty}^{+\infty} |\Delta(u)| d\alpha(u)$ est alors une vraie variable aléatoire".

Ainsi, ici encore $\Delta^*(\varphi)$ peut s'écrire $\langle g, \Delta \rangle$, où g appartient à l_α^∞ et Δ à l_α^1 , et donc peut être considérée comme une fonctionnelle linéaire sur l_α^∞ .

D - Δ^* FONCTIONNELLE LINEAIRE SUR D_α^p

1 - Désignons enfin par :

l_α^p l'espace des fonctions intégrables par rapport à α sur tout compact et telles que :

$$\int_{-\infty}^{+\infty} |g|^p d\alpha < \infty,$$

et par :

D_α^p l'espace des fonctions φ s'écrivant :

$$\int_a^x g(u) d\alpha(u)$$

où $g(u)$ appartient à l_α^p .

Lemme - "Si $\alpha(u)$ est une fonction numérique non-décroissante telle que :

$$\alpha^2(u) \{ [1 - F(x)] F(-x) \}^{q-1} \in C_{p_X},$$

et si :

$$g(u) \in l_\alpha^p$$

avec :

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (p > 1),$$

alors $\int_a^x g(u) d\alpha(u)$ appartient à $C_{p_x}^2$."

Supposons que $\alpha(a) = 0$, on a, si $\varphi \in D_a^p$:

$$\int_{-\infty}^{+\infty} \varphi^2 dP_x = \int_{u,v \in (a,x)} g(u) g(v) d\alpha(u) d\alpha(v) dP_x$$

posons :

$$F_x(u, v) = P_x \{ x : u, v \in (a, x) \}$$

on a :

$$\int_{-\infty}^{+\infty} \varphi^2 dP_x < \left[\int [g(u) g(v)]^p d\alpha(u) d\alpha(v) \right]^{\frac{1}{p}} \left[\int F_x^q(u, v) d\alpha(u) d\alpha(v) \right]^{\frac{1}{q}}$$

Donc si :

$$J = \int F_x^q(u, v) d\alpha(u) d\alpha(v) < \infty,$$

alors φ appartient à $C_{p_x}^2$. Or on peut écrire J :

$$J = \frac{1}{q} \int_{a < u, v < x} [1 - F_x(x)]^{q-1} d\alpha(u) d\alpha(v) dP_x + \frac{1}{q} \int_{x < u, v < a} [F_x(x)]^{q-1} d\alpha(u) d\alpha(v) dP_x$$

soit finalement :

$$qJ = \int_a^\infty [1 - F_x(x)]^{q-1} \alpha^2(x) dP_x + \int_{-\infty}^a [F_x(x)]^{q-1} \alpha^2(x) dP_x$$

et sous cette forme on reconnaît aisément l'énoncé du lemme : remarquons d'ailleurs que la condition : $\alpha \in C_{p_x}^2$ est une condition suffisante.

2 - La conclusion de ce paragraphe sera comme précédemment :

Proposition III - "Il existe une fonction aléatoire $\Delta(u)$ de l_a^q qui est représentée dans $L_{p_x}^2$ par la série (1)".

Comme il paraît vraisemblable que la même fonction $\Delta(u)$ satisfiera aux trois propositions à démontrer et que cette fonction représentera donc finalement la loi d'erreur, comme fonctionnelle linéaire aussi bien que dans $L^2_{P_X}$, on l'appellera : fonction "déterminante" de la loi d'erreur.

E - FONCTION DETERMINANTE DE LA LOI D'ERREUR

1 - Par son origine même cette fonction $\Delta(u)$ sera la limite en loi de la fonction aléatoire $\sqrt{n} [F_n(x) - F_X(x)]$ quand $n \rightarrow +\infty$ et devra avoir les propriétés suivantes :

(A) presque sûrement continue et bornée,

(B) si α satisfait aux conditions du lemme (C - 1), $\Delta^\alpha(u)$ est presque sûrement intégrable par rapport à α . En particulier, pour $q = 2$, on en déduit que $\Delta(u)$ appartient presque sûrement à L^2_β et y est représentée par la série (1) ; cette remarque peut permettre de simplifier la démonstration du théorème suivant, théorème fondamental pour l'étude et la définition de "l'erreur statistique".

2 - Théorème de Prohorov :

"Soit C_0 l'espace des fonctions numériques continues et bornées. Quand $n \rightarrow \infty$, la loi de la fonction aléatoire :

$$\sqrt{n} [F_n(x) - F_X(x)]$$

tend vers une mesure de probabilité sur C_0 ; cette mesure est la loi d'une fonction aléatoire $\Delta(x)$ de C_0 qui satisfait de plus à la condition (B)".

La démonstration donnée par Prohorov dans [6] est faite dans le cas d'une répartition P_X uniforme sur le segment $(0,1)$ et procède du plan suivant :

Soit $\mathcal{S}$ l'ensemble des fonctions $f(t)$, réelles, sur $(0,1)$ qui en chaque point ont :

une limite à droite $f^+(t)$

une limite à gauche $f^-(t)$

telles que :

$$f(t) = f^+(t) \quad \text{ou} \quad f^-(t)$$

l'égalité pouvant dépendre de t ; on ajoutera :

$$f^+(0) = f^-(1) = 0.$$

Soit $\tilde{\mathcal{F}}$ l'espace des classes d'équivalence sur $\mathcal{F}$ obtenues en identifiant deux fonctions f_1 et f_2 telles que :

$$f_1^+(t) = f_2^+(t) \quad \text{et} \quad f_1^-(t) = f_2^-(t).$$

On peut définir sur $\tilde{\mathcal{F}}$ une distance $d(\varphi, \varphi')$, telle que $\tilde{\mathcal{F}}$ muni de cette distance soit séparable et complet ; de plus, l'ensemble des fonctions continues est d -fermé dans $\tilde{\mathcal{F}}$, la d -convergence y étant équivalente à la convergence uniforme.

Soit $\mathcal{M}(\tilde{\mathcal{F}})$ l'espace des mesures sur $\tilde{\mathcal{F}}$; on montre que M_n , loi de probabilité de $\sqrt{n}[F_n(x) - F_x(x)]$, tend faiblement vers M , loi de probabilité de $\Delta_0(x)$, fonction aléatoire laplacienne, presque sûrement continue, de covariance :

$$\inf(x, y) - xy.$$

La convergence faible est dans $\mathcal{M}(\tilde{\mathcal{F}})$ définie par l'une des trois conditions équivalentes :

(C₁) Quelle que soit ϑ , fonctionnelle continue sur $\tilde{\mathcal{F}}$:

$$E\{\vartheta[\sqrt{n}(F_n(x) - F_x(x))]\} \longrightarrow E\{\vartheta(\Delta_0(x))\}$$

(C₂) $L(M_n, M) \longrightarrow 0$, L étant la distance de Lévy.

(C₃) Quelle que soit ϑ , fonctionnelle continue M - presque partout sur $\tilde{\mathcal{F}}$, la loi de la variable aléatoire :

$$\vartheta[\sqrt{n}(F_n(x) - F_x(x))]$$

tend vers celle de la variable aléatoire :

$$\vartheta(\Delta_0(x)).$$

3 - De cette dernière expression on déduit des corollaires, qui d'ailleurs ont été pour la plupart établis directement.

"Théorème de Kolmogorov-Smirnov".

Comme la distance d est équivalente, pour les fonctions continues de $\tilde{\mathcal{F}}$ à :

$$\sup_x |f(x) - g(x)|$$

on prend dans C_3 , pour fonctionnelle Φ , l'application :

$$\varphi \in \tilde{\mathcal{F}} \longrightarrow \sup |\varphi|.$$

On peut, sans changer la loi limite, rendre $F_n(x)$ continue en interpolant linéairement, puisque les sauts ont tous pour valeurs $\frac{1}{\sqrt{n}}$.

"Théorème de Von-Mises".

On peut même l'énoncer ici sous la forme :

Si $\alpha(x)$ satisfait aux conditions du lemme (C - 1), la loi de la variable aléatoire :

$$\int_{-\infty}^{+\infty} |\sqrt{n} (F_n(x) - F_x(x))|^q d\alpha(x)$$

tend vers celle de la variable aléatoire :

$$\int_{-\infty}^{+\infty} \Delta_o^q(x) d\alpha(x).$$

4 - Forme réduite de la fonction déterminante d'erreur.

Depuis le début de ce travail, il n'a volontairement pas été question du changement $y = F_x(x)$, pourtant si souvent utilisé en Calcul des Probabilités.

Plus précisément, soit la transformation du plan qui au point $P(x, y)$ fait correspondre le point $Q[F_x(x), y]$; soit $\tilde{\Delta}(u)$ la transformée par cette opération du graphe de $\Delta(u)$.

Si $F_x(x)$ est continue et strictement croissante, $\tilde{\Delta}(u)$ est le graphe de la fonction $\Delta[F_x^{-1}(x)]$, définie sur $(0, 1)$. Cette fonction aléatoire a une loi qui ne dépend plus de F_x et peut alors s'écrire :

$$\tilde{\Delta}(u) = X(t) - tX(1) \quad 0 \leq t \leq 1.$$

où $X(t)$ est sur le segment $(0, 1)$, la fonction aléatoire réduite de

Wiener-Lévy. La formule (1) de la page 18 donne la représentation classique dans $L^2_{(0,1)}$, de cette fonction aléatoire à partir de sa covariance ; c'est-à-dire que l'on peut choisir $\{f_k\}$ telle que $\int_0^1 f_k(x) dx$ soit encore ortho-normée et l'on a :

$$\tilde{\Delta}(u) = \sum_{k=1}^{\infty} U_k \frac{\sqrt{2}}{k\pi} \cdot \sin k\pi t$$

dont on déduit le théorème de Von-Mises :

$$\int_0^1 \tilde{\Delta}^2(u) du = w^2 \text{ a pour fonction caractéristique :}$$

$$\sqrt{\frac{\sqrt{2it}}{\sin \sqrt{2it}}}$$

Si $F_x(x)$ n'est pas continue, $\tilde{\Delta}(u)$ n'est définie que sur J , image de la droite par l'application $x \rightarrow F_x(x)$; il est intéressant de noter qu'alors :

$$\int \Delta(x)^2 dP_x = \int_J \tilde{\Delta}(u)^2 du$$

et :

$$\sup_x |\Delta(x)| = \sup_J |\tilde{\Delta}(u)|$$

sont des variables aléatoires dont les lois sont respectivement moins étalées que celles de Von-Mises et de Kolmogoroff-Smirnov ; ces théorèmes donnent donc seulement une inégalité, intéressante d'ailleurs ; par exemple :

$$\forall z > 0 \quad \lim_{n \rightarrow \infty} \text{Prob.} \{ \sqrt{n} |F_n(x) - F_x(x)| < z \} < L(z).$$

F - CONCLUSION : PROPRIETES DE P_n COMME ESTIMATEUR DE P_x

Ces deux chapitres ont mis en évidence un ensemble de propriétés, définissant l'estimation de P_x par P_n ; en voici le résumé, auquel on se réfèrera dans toute la suite :

Convergence simple.

(C₁) $P_n^*(\varphi) \longrightarrow P_x^*(\varphi)$ p.s., si et seulement si $P_x^*(\varphi)$ existe

(C₂) $P_n^*(\varphi) \longrightarrow \tilde{P}_x^*(\varphi)$ p., si et seulement si $\tilde{P}_x^*(\varphi)$ existe

(C₃) $P_n^*(\varphi) \longrightarrow P_x^*(\varphi)$ p.c.s., si et seulement si $P_x^*(\varphi^2)$ existe.

la convergence, uniforme sur des sous-ensembles de D_a , est entièrement déduite des deux théorèmes :

(C₄) $\sup_x |F_n(x) - F_x(x)| \longrightarrow 0$ p.c.s.

(C₅) si $\alpha \in C_{p_x}$, $\int_{-\infty}^{+\infty} |F_n(x) - F_x(x)|^r d\alpha(x) \longrightarrow 0$ p.s.

Loi d'erreur.

$\sqrt{n}[P_n^* - P_x^*]$ a, quand $n \longrightarrow \infty$, une limite en loi, application linéaire aléatoire de $L_{p_x}^2$ dans R , dont les restrictions à des sous espaces de D_a sont des fonctionnelles linéaires. La loi de probabilité de tous ces éléments aléatoires est entièrement déterminée par celle de la fonction aléatoire laplacienne, limite en loi de la fonction aléatoire $\sqrt{n}[F_n(x) - F_x(x)]$.

CHAPITRE III

PRODUITS DE RÉPARTITIONS EMPIRIQUES

L'objet des chapitres suivants est d'examiner le problème que voici : (cf. Introduction).

Soient P_n et Q_m deux répartitions empiriques, identiques ou non, estimant respectivement P_x et P_y : comment $P_n \times Q_m$ estime-t-il $P_x \times P_y$ dans R^2 ?

Soit $\varphi(x, y)$ une fonction borélienne dans R^2 ; la loi de probabilité $P_{x,y}$ du vecteur (X, Y) fait correspondre à φ , d'une part un nombre :

$$\int \varphi dP_{x,y}$$

d'autre part une loi de probabilité P^φ sur R : la loi de la variable aléatoire $\varphi(X, Y)$. Ces deux points de vue sont équivalents, mais entraînent des problèmes d'estimation différents; on établira dans les deux cas des propriétés analogues à celles de la conclusion du Chapitre II.

Pour des raisons de clarté, la démonstration de certains théorèmes nouveaux sera donnée dans des chapitres ultérieurs; on distinguera enfin le cas où P_n et Q_m sont des répartitions empiriques indépendantes, du cas où P_n et Q_m proviennent du même échantillon.

A - CONVERGENCE SIMPLE DE $(P_n \times Q_m)^*$ VERS $(P_x \times P_y)^*$

On établira les propositions :

Proposition P_1 - " $\{X_i\}$ $i = 1, \dots, n$ et $\{Y_j\}$ $j = 1, \dots, m$ étant deux échantillons indépendants, respectivement des variables aléatoires X et Y on a :

$$\text{Si} \quad \varphi(x, y) \in C_{P_X \times P_Y},$$

$$\frac{1}{mn} \sum_{i,j} \varphi(X_i, Y_j) = (P_n \times Q_m)^* (\varphi)$$

$$\text{et} \quad (P_n \times Q_m)^* (\varphi) \longrightarrow (P_X \times P_Y)^* (\varphi) \quad \text{p. s.}$$

quand m et n tendent indépendamment vers l'infini".

Proposition P' - " $\{X_i\}$ $i = 1, \dots, n$ étant un échantillon de la variable aléatoire X on a :

$$\text{Si} \quad \varphi(x, y) \in C_{P_X \times P_Y},$$

$$\frac{1}{n(n-1)} \sum_{i \neq j} \varphi(X_i, X_j) \longrightarrow (P_X \times P_Y)^* (\varphi) \quad \text{p. s.}$$

quand $n \rightarrow \infty$ ".

D'une façon analogue, on aura convergence en probabilité ou presque complètement sûrement, suivant que l'on aura :

$$\varphi \in C_{P_X \times P_Y}^0 \quad \text{ou} \quad \varphi \in C_{P_X \times P_Y}^2.$$

On notera que si, φ , étant une fonction borélienne fixée et donnée à l'avance, on étudie comment

$$(P_n \times Q_m)^{q_0} \quad \text{estime} \quad (P_X \times P_Y)^{q_0}$$

on est conduit à démontrer les mêmes théorèmes pour des fonctions boréliennes de la forme $f(\varphi)$, f étant une application borélienne de R dans R .

Ces théorèmes de loi des grands nombres pour des variables aléatoires qui ne sont pas indépendantes seront établis aux chapitres IV et V.

B - CONVERGENCE UNIFORME DE $(P_n \times Q_m)^*$ VERS $(P_X \times P_Y)^*$

1 - La fonction de répartition empirique associée à la loi de probabilité empirique $P_n \times Q_m$ est égale à

$$F_n(x) G_m(y)$$

et l'on peut écrire :

$$|F_n(x)G_m(y) - F_x(x)F_y(y)| < F_x(x)|G_m(y) - F_y(y)| + F_y(y)|F_n(x) - F_x(x)| \\ + |F_n(x) - F_x(x)| \cdot |G_m(y) - F_y(y)|,$$

donc :

$$\sup_{x,y} |F_n(x) G_m(y) - F_x(x) F_y(y)| \longrightarrow 0 \quad \text{p. c. s.}$$

quand m et n tendent indépendamment vers l'infini.

Mais il faut établir de plus que, φ_0 étant encore ici fixée, la répartition empirique $K_{m,n}(z)$ de l'ensemble des variables aléatoires $\varphi_0(X_i, Y_j)$, tend p. c. s. uniformément vers $K(z)$ fonction de répartition de la variable aléatoire $Z = \varphi_0(X, Y)$.

On se reportera à la démonstration donnée dans [8] du théorème de Glivenko-Cantelli, pour noter que celle-ci ne fait pas intervenir l'indépendance des variables aléatoires de l'échantillon, mais suppose seulement qu'elles satisfont à la loi forte des grands nombres. On a finalement :

Proposition - "Quelles que soient les fonctions boréliennes $\varphi_0(x, y)$ et $\varphi_1(x, y)$

$$\sup_{\lambda, \mu} \left| \iint_{\varphi_0(x,y) < \lambda, \varphi_1(x,y) < \mu} d[P_n \times Q_m - P_x \times P_y] \right| \longrightarrow 0 \quad \text{p. c. s.}$$

Si m et n tendent vers l'infini indépendamment".

2 - La démonstration de la page 10 ne faisant intervenir elle aussi, que la loi des grands nombres et le théorème précédent, on en déduit que si $\alpha(z)$ est une fonction borélienne non décroissante de C_{P_2} , on a

$$\int |K_{m,n}(z) - K(z)|^r d\alpha(z) \longrightarrow 0 \quad \text{p. s.}$$

quand m et n tendent vers l'infini simultanément.

On établira aussi que

$$\iint |F_n(x) G_m(y) - F_x(x) F_y(y)|^r d\alpha(x, y) \longrightarrow 0 \quad \text{p. s.}$$

si $\alpha(x, y)$ n'est pas décroissante et appartient à $C_{P_x \times P_y}$.

On écrit en effet cette intégrale dans le rectangle $|x| < A, |y| < B$, où la convergence vers 0 découle du théorème précédent; en dehors

du rectangle la loi des grands nombres rend, presque sûrement, l'intégrale aussi petite que l'on veut.

C - LOI D'ERREUR

On doit faire ici des hypothèses sur la façon dont m et n tendent vers l'infini; on supposera que :

$$(C_1) \quad \frac{m}{m+n} \rightarrow a \quad \text{et} \quad \frac{n}{m+n} \rightarrow b \quad (a+b=1)$$

Il est alors clair qu'on a sur $C_{P_X \times P_Y}^2$ une application aléatoire entièrement définie par la fonction déterminante d'erreur, fonction aléatoire à deux variables qui peut se représenter par :

$$\lambda F_X(\mathbf{x}) \Delta_Y(\mathbf{y}) + \mu F_Y(\mathbf{y}) \Delta_X(\mathbf{x})$$

où λ et μ sont des constantes, Δ_X et Δ_Y les fonctions déterminantes d'erreur respectivement des lois P_X et P_Y . On établira le théorème central limite suivant :

Proposition [1] - "Si $\varphi \in C_{P_X \times P_Y}^2$, quand m et n tendent vers l'infini suivant C_1 , alors

$$\mathcal{L} \left\{ \sqrt{m+n} \left[\frac{1}{mn} \sum_{i,j} \varphi(X_i, Y_j) - \int \varphi dP_X dP_Y \right] \right\} \rightarrow N(0, \sigma). "$$

Si l'on se restreint aux fonctions φ de la forme $f(\varphi_0)$, on obtient la loi d'erreur associée à la répartition empirique $K_{m,n}(z)$ et cette loi d'erreur possède une propriété remarquable qui sera établie au chapitre V.

D - CONCLUSION

On vient de montrer que des variables aléatoires non indépendantes et non en chaîne strictement stationnaire, peuvent satisfaire aux mêmes propriétés qu'un échantillon en ce qui concerne l'estimation de la loi commune à ces variables. Les deux propriétés fondamentales qui restent et qui entraînent d'ailleurs toutes les autres propriétés connues d'une répartition empirique classique, sont donc :

(A) $X_1, \dots, X_n$ étant des variables aléatoires de même loi P_X

$\frac{\varphi(X_1) + \dots + \varphi(X_n)}{n}$ converge sous les mêmes conditions qu'une suite de variables aléatoires indépendantes.

(B) L'application linéaire aléatoire représentant la loi d'erreur admet une covariance qui, comme noyau dans $L^2_{\mathcal{P}_X}$, est inférieure à la covariance associée à un échantillon.

CHAPITRE IV

LOIS DES GRANDS NOMBRES POUR DES VARIABLES ALÉATOIRES INTER-DÉPENDANTES

Nous établissons ici les propositions P_1 et P'_1 , annoncées au chapitre précédent, paragraphe A : soient donc $X, X_1, \dots, X_n$ des variables aléatoires de loi P_x , $Y, Y_1, \dots, Y_m$ des variables aléatoires de loi P_y , toutes ces variables aléatoires étant indépendantes entre elles; soient enfin $H(x, y)$ une fonction borélienne et P_z la loi de la variable aléatoire

$$Z = H(X, Y)$$

$K(z)$ désignant la fonction de répartition de P_z .

En plus de ce que nous avons annoncé, on pourra noter l'utilisation des suites doubles de variables aléatoires, qui conduit à introduire de nouveaux modes de convergence, en probabilité ou presque sûrement. Enfin nous considérerons successivement les propositions du type P , puis celles du type P' .

A - VARIABLES ALEATOIRES ENCHAINÉES, TYPE A

1 - Loi faible.

$$''\text{Si} : (C_1) \quad n[1 - K(+n) + K(-n)] \longrightarrow 0, \quad \text{quand } n \longrightarrow \infty$$

$$\text{et :} \quad \int_{-u}^{+u} z \, dK(z) \longrightarrow a \quad \text{quand } u \longrightarrow \infty$$

$(P_n \times Q_m)^*(H) \longrightarrow a$, en probabilité quand m et n tendent vers l'infini simultanément''.

Démonstration -

Soient :

$$mn = v$$

$$\text{et} \quad H_\nu(\mathbf{x}, \mathbf{y}) = \begin{cases} H(\mathbf{x}, \mathbf{y}) & \text{si} \quad |H(\mathbf{x}, \mathbf{y})| \leq \nu \\ 0 & \text{si} \quad |H(\mathbf{x}, \mathbf{y})| > \nu \end{cases}$$

la condition (C_1) entraîne :

$$(P_n \times Q_m)^*(H) - (P_n \times Q_m)^*(H_\nu) \longrightarrow 0 \text{ en probabilité.}$$

Posons maintenant :

$$a_\nu = \int_{-\nu}^{+\nu} z \, dK(z)$$

et

$$V(\mathbf{x}, \mathbf{y}) = H_\nu(\mathbf{x}, \mathbf{y}) - a_\nu$$

$$V_1(\mathbf{x}) = E\{V(\mathbf{x}, \mathbf{Y})\}$$

$$V_2(\mathbf{y}) = E\{V(\mathbf{X}, \mathbf{y})\}.$$

Soit

$$W(\mathbf{x}, \mathbf{y}) = V(\mathbf{x}, \mathbf{y}) - V_1(\mathbf{x}) - V_2(\mathbf{y})$$

on a :

$$E\{W(\mathbf{X}, \mathbf{Y}) | \mathbf{Y}\} = E\{W(\mathbf{X}, \mathbf{Y}) | \mathbf{X}\} = E\{W(\mathbf{X}, \mathbf{Y})\} = 0$$

d'où

$$E\left\{\sum_{i,j} W(\mathbf{X}_i, \mathbf{Y}_j)\right\}^2 = \nu \, E\{W(\mathbf{X}, \mathbf{Y})^2\}.$$

Or

$$E\{W(\mathbf{X}, \mathbf{Y})^2\} = E\{V(\mathbf{X}, \mathbf{Y})^2\} - E\{V_1(\mathbf{X})^2\} - E\{V_2(\mathbf{Y})^2\}$$

et donc

$$E\{W(\mathbf{X}, \mathbf{Y})^2\} \leq \int_{-\nu}^{+\nu} z^2 \, dK(z) - a_\nu^2.$$

Ecrivons alors :

$$(P_n \times Q_m)^*(H) - a_\nu = \frac{1}{\nu} \sum_{i,j} W(\mathbf{X}_i, \mathbf{Y}_j) + \frac{1}{n} \sum_i V_1(\mathbf{X}_i) + \frac{1}{m} \sum_j V_2(\mathbf{Y}_j)$$

on voit ainsi que :

$$\text{d'une part} \quad E\{V_1(\mathbf{X})\} = E\{V_2(\mathbf{Y})\} = 0 \quad \text{entraîne que}$$

$$\frac{1}{n} \sum_i V_1(X_i) \quad \text{et} \quad \frac{1}{m} \sum_j V_2(Y_j)$$

tendent presque sûrement vers 0 quand m et n tendent vers l'infini.

d'autre part, (C_1) entraînant que

$$\frac{1}{v} \int_{-v}^{+v} z^2 dK(z) \longrightarrow 0 \quad \text{si} \quad v \longrightarrow +\infty,$$

$$\frac{1}{v} \sum_{i,j} W(X_i, Y_j) \longrightarrow 0 \quad \text{en probabilité.}$$

On remarquera que, si l'on retient seulement la condition C_1 on a, comme dans le théorème classique, la stabilité par rapport à la suite des a_v .

2 - Lois fortes.

Ayant ici des suites doubles de variables aléatoires, on peut se placer, pour définir la convergence presque sûre, à deux points de vue :

a) n et m sont fonctions d'un paramètre p , entier par exemple, et satisfont à la condition :

$$(C') \quad p > q \implies m(p) > m(q) \quad \text{ou} \quad n(p) > n(q);$$

ce qui est équivalent à dire que l'indication par le paramètre p des points $M(m, n)$ parcourus dans le plan (m, n) n'est jamais décroissante pour l'ordre produit; et d'autre part la série $\frac{1}{m(p) n(p)}$ est supposée convergente.

b) n et m sont indépendants, c'est-à-dire que l'on devra considérer la réalisation simultanée de toutes les inégalités :

$$|S(m, n)| < \varepsilon \quad \text{pour} \quad m \text{ et } n > N.$$

Dans les deux cas nous suivrons le schéma de démonstration classique, en établissant les lemmes qui suivent.

3 - Lemme 1.

"Soient deux suites de variables aléatoires indépendantes $\{X_i\}$,

$\{Y_i\}$; à tout couple (X_i, Y_j) correspond fonctionnellement la variable aléatoire V_j^i telle que :

$$(C_2) \quad E\{V_j^i\} = E\{V_j^i | Y_j\} = E\{V_j^i | X_i\} = 0.$$

Si l'on pose alors

$$S_p = \sum_{1 \leq i, j \leq p} V_j^i$$

$$\text{Prob.} \left\{ \sup_{p=1, \dots, n} |S_p| > \varepsilon \right\} \leq \frac{1}{\varepsilon^2} E(S_n^2),$$

avec

$$E(S_n^2) = \sum_{1 \leq i, j \leq n} (\sigma_j^i)^2 \quad \text{si} \quad (\sigma_j^i)^2 = E\{(V_j^i)^2\}."$$

Démonstration.

Soit E_k l'évènement suivant :

Le premier S_p de module supérieur à ε est S_k ; on a

$$E(S_n^2) = \sum_k P(E_k) E(S_n^2 | E_k).$$

Ecrivons

$$S_n = S_k + S_{n,k}^i$$

où

$$S_{n,k}^i = \sum_I V_j^i$$

I étant l'ensemble d'indices défini par :

$$i, j \leq n, \quad i \text{ ou } j \geq k;$$

on a alors

$$E\{V_j^i V_{j'}^{i'} | E_k\} = 0$$

$$\text{si} \quad i, j \leq k \quad \text{et} \quad (i', j') \in I,$$

Ceci en vertu de C_2 et de ce que E_k ne concerne que les X_i et Y_j tel que i et j soient inférieurs ou égaux à k . Il vient finalement

$$E\{S_n^2 | E_k\} \geq \varepsilon^2$$

et donc la proposition puisque l'on cherchait à majorer :

$$\sum_k P(E_k).$$

Corollaire 1 - Les hypothèses étant les mêmes, soit

$$S(m, n) = \sum_{i \leq n, j \leq m} V_j^i;$$

si m et n sont fonctions d'un paramètre p et satisfont à C' , la démonstration précédente est encore valable et l'on a :

$$\text{Prob.} \left\{ \sup_{p=1, \dots, p_0} |S(n_p, m_p)| \geq \varepsilon \right\} \leq \frac{1}{\varepsilon^2} E \{ S(p_0)^2 \}.$$

Corollaire 2 - Si maintenant m et n sont indépendants on en déduit :

$$\text{Prob.} \left\{ \sup_{n \leq n_0, m \leq m_0} |S(n, m)| > \varepsilon \right\} \leq \frac{\inf(n_0, m_0)}{\varepsilon^2} E \{ S(n_0, m_0)^2 \}.$$

En effet, l'ensemble d'indices (n, m) :

$$n \leq n_0, \quad m \leq m_0 \quad (m_0 \leq n_0)$$

est la réunion des familles d'indices I_q

$$I_q = \begin{cases} n = 1, & m \leq q \\ n \leq n_0, & m = q \\ n = n_0, & q < m \leq m_0; \end{cases}$$

on a ainsi m_0 familles d'indices, indicées par q de façon à satisfaire à C' , l'inégalité de Boole et le corollaire 1 établissent alors le corollaire 2.

4 - Critère 1.

"Si $\frac{(\sigma_j^1)^2}{[\text{Sup}(i, j)]^4}$ est le terme général d'une série double convergente, alors $\frac{1}{p^2} S_p$ admet une limite presque sûre."

Les notations sont les mêmes qu'au paragraphe 3, et l'on peut écrire la suite d'expressions où chacune d'elles majore la précédente :

$$\begin{aligned}
& 1 - \text{Prob.} \left\{ \frac{1}{p^2} |S_p| < \varepsilon, \quad \forall p \geq p_0 \right\} \\
& \sum_{2^\mu > p_0} \text{Prob.} \{ \text{Sup} |S_p| > \varepsilon \cdot p^2, \quad 2^\mu \leq p < 2^{\mu+1} \} \\
& \sum_{2^\mu > p_0} \text{Prob.} \{ \text{Sup} |S_p| > \varepsilon \cdot 4^\mu, \quad 2^\mu \leq p < 2^{\mu+1} \} \\
& \frac{1}{\varepsilon^2} \sum_{2^\mu > p_0} \left\{ \frac{1}{16^\mu} \sum_{i, j \leq 2^{\mu+1}} (\sigma_j^i)^2 \right\}
\end{aligned}$$

Soit en intervertissant les sommations

$$\frac{64 \log 2}{\varepsilon^2} \left\{ \sum_{i \text{ ou } j > p_0} \frac{(\sigma_j^i)^2}{[\text{Sup}(i, j)]^4} + \frac{1}{p_0^4} \sum_{i, j < p_0} (\sigma_j^i)^2 \right\};$$

Cette dernière expression tend vers 0, quand $p_0 \rightarrow \infty$, ce qui établit le critère.

5 - Comme dans la démonstration de la loi forte des grands nombres, on peut établir le résultat précédent, d'une façon différente, en montrant les propositions suivantes :

Supposons d'une façon plus générale que m et n dépendent du paramètre p en satisfaisant à la condition (C') et posons

$$S'_p = S(m(p), n(p)),$$

on a :

Critère 1 bis - "Si $(\sigma_j^i)^2$ est le terme général d'une série double convergente, S'_p admet presque sûrement une limite finie quand $p \rightarrow \infty$."

Soient les nouvelles variables aléatoires déduites des V_j^i par :

$$W_j^i = \begin{cases} V_j^i & \text{si } i > n(p_0) \text{ ou } j > m(p_0) \\ 0 & \text{si } i \leq n(p_0) \text{ et } j \leq m(p_0) \end{cases}$$

Ces variables aléatoires W_j^i remplissent la condition C_2 et l'on a

$$T(p) = \sum_{i \leq n(p), j \leq m(p)} W_j^i = S'_p - S_{p_0} \quad \text{si } p > p_0$$

et

$$T(p) = 0 \quad \text{si } p \leq p_0.$$

Appliquons le corollaire 1 aux T_p ; il vient :

$$\text{Prob.}\{\text{Sup}|T(p)| > \varepsilon, p \leq p_1\} = \text{Prob.}\{\text{Sup}|S'_p - S'_{p_0}| > \varepsilon, p_0 < p \leq p_1\}$$

et ces probabilités sont majorées par

$$\frac{1}{\varepsilon^2} E\{T(p_1)^2\} = \frac{1}{\varepsilon^2} \sum_I (\sigma_j^1)^2,$$

où I est l'ensemble d'indices défini par :

$$\begin{cases} i \leq n(p_1) & \text{et} & j \leq m(p_1) \\ i > n(p_0) & \text{ou} & j > m(p_0); \end{cases}$$

Si l'on fait tendre p_1 puis p_0 vers l'infini, on a

$$\sum_I (\sigma_j^1)^2 \longrightarrow 0$$

et le critère de Cauchy donne la proposition pour la suite S'_p .

Corollaire 1 bis - "Si $\frac{(\sigma_j^1)^2}{g^2(i, j)}$ est le terme général d'une série double convergente on a

$$\frac{1}{f(p)} S'_p \longrightarrow 0 \quad \text{p. s., quand } p \longrightarrow \infty$$

la correspondance entre g et f est définie par :

à tout couple (i, j) on fait correspondre un nombre k , unique tel que :

$$\begin{aligned} i &\leq n(k), & j &\leq m(k) \\ i &> n(k-1) & \text{ou} & j > m(k-1), \end{aligned}$$

alors

$$g(i, j) = f(k(i, j)).$$

En particulier $k(i, j) = \sup(i, j)$ quand

$$m = n = p \quad ''.$$

Posons

$$U_k = \sum_{I_k} V_j^i,$$

I_k étant l'ensemble d'indices :

$$\begin{cases} i \leq n(k), & j \leq m(k) \\ j > n(k-1) & \text{ou} & i > n(k-1); \end{cases}$$

d'après C' , I_k n'est pas vide et à un couple (i, j) correspond un seul k .

D'après le lemme de Kronecker, si $\frac{U_k}{f(k)}$ est le terme général d'une série convergente on a :

$$\frac{1}{f(p)} \sum_{k=1}^p U_k = \frac{1}{f(p)} S_p' \rightarrow 0$$

quand $p \rightarrow \infty$; or

$$\sum_1^p \frac{U_k}{f(k)} = \sum_{i \leq n(p), j \leq m(p)} \frac{V_j^i}{g(i, j)}$$

et il suffit d'appliquer le critère 1 bis aux variables $\frac{V_j^i}{g(i, j)}$ qui remplissent bien les conditions C_2 . Pour S_p et $f(p) = p^2$ on retrouve le critère 1.

6 - Revenons au cas où m et n tendent indépendamment vers l'infini et établissons le critère 2 :

"Si $\frac{(\sigma_j^i)^2}{[\sup(i, j)]^2 \inf(i, j)}$ est le terme général d'une série double convergente, $\frac{1}{mn} S(m, n) \rightarrow 0$ p.s., quand m et n tendent vers l'infini indépendamment".

Démonstration - On a :

$$\text{Prob.} \left\{ \sup_I |S(n, m)| > \varepsilon, \quad I : n_1 \leq n \leq n_2, \quad m_1 \leq m \leq m_2 \right\}$$

$$\leq \frac{\inf\{(n_2 - n_1), (m_2 - m_1)\}}{\varepsilon^2} \sum_{1 \leq n_2, j \leq m_2} (\sigma_j^i)^2$$

Or, si l'on pose

$$P(\lambda, \mu) = \text{Prob.} \left\{ \text{Sup} \left| \frac{S(m, n)}{m n} \right| > \varepsilon, \quad 2^\lambda < n \leq 2^{\lambda+1}, \quad 2^\mu < m \leq 2^{\mu+1} \right\}$$

on a :

$$1 - \text{Prob.} \{ |S(m, n)| < \varepsilon . m . n, \quad \forall m \text{ et } n > N \} \quad (2)$$

est inférieure à

$$\sum_{2^\lambda, 2^\mu > N} P(\lambda, \mu). \quad (1)$$

Mais on a :

$$P(\lambda, \mu) \leq \text{Prob.} \{ |S(m, n)| < \varepsilon 2^{\lambda+\mu}, \quad 2^\lambda < n \leq 2^{\lambda+1}, \quad 2^\mu < m \leq 2^{\mu+1} \}$$

Soit encore

$$P(\lambda, \mu) \leq \frac{2^{\inf(\lambda, \mu)}}{\varepsilon^2 4^{\lambda+\mu}} \sum_{1 \leq 2^{\lambda+1}, j \leq 2^{\mu+1}} (\sigma_j^1)^2;$$

en reportant cette dernière majoration dans (1), il vient à des constantes numériques près, comme ultime majoration de (2) :

$$\sum_{i, j} \frac{(\sigma_j^1)^2}{[\text{Sup}(2N, i, j)]^3} + \sum_{i \text{ ou } j > N} \frac{(\sigma_j^1)^2}{[\text{Sup}(i, j)]^2 \text{Sup}[N, \inf(i, j)]}.$$

Si $\frac{(\sigma_j^1)^2}{[\text{Sup}(i, j)]^3}$ est le terme général d'une série convergente, on voit par définition et par application du lemme de Kronecker que la première somme tend vers 0 avec $\frac{1}{N}$. D'autre part, la seconde somme est inférieure à

$$\sum_{i \text{ ou } j > N} \frac{(\sigma_j^1)^2}{[\text{Sup}(i, j)]^2 \inf(i, j)}$$

expression qui tend vers 0 avec $\frac{1}{N}$ par hypothèse.

Nous sommes maintenant en état de démontrer les théorèmes de convergence forte que nous avons en vue.

7 - Théorème I.

"Si $E(Z)$ existe, quelles que soient les fonctions $n(p)$ et $m(p)$

satisfaisant à C' , on a

$$(P_n \times Q_m)^* (H) \longrightarrow (P_x \times P_y)^* (H) = E(Z), \quad p.s.,$$

quand $p \longrightarrow +\infty$.

Démonstration - Revenons aux notations du corollaire 1 bis, avec ici :

$$f(p) = m(p) n(p)$$

Si I_p désigne l'ensemble d'indices (i, j) défini par

$$i \leq n(p) \quad \text{et} \quad j \leq m(p),$$

il vient :

$$k(i, j) = p \quad \text{si} \quad (i, j) \in I_p \cap I_{p-1};$$

$k(i, j)$ est une fonction bien définie parce que les fonctions $n(p)$ et $m(p)$ satisfont à C' , et l'on a

$$g(i, j) = m[k(i, j)] n[k(i, j)].$$

Lemme - "Si $u(n)$ est le terme général d'une série convergente et si $u(n)$ est positive et décroissante à partir d'un certain rang,

$$v(i, j) = u[g(i, j)]$$

est le terme général d'une série double convergente".

On a en effet :

$$\sum_n u(n) = \sum_k \left\{ \sum_{m(k-1)n(k-1) < n \leq m(k)n(k)} u(n) \right\}$$

est supérieure à

$$\sum_k [m(k) n(k) - m(k-1) n(k-1)] u(m(k) n(k))$$

Soit encore

$$\sum_k \left\{ \sum_{(i,j) \in I_k \cap I_{k-1}} v(i, j) \right\} = \sum_{i,j} v(i, j)$$

Soient maintenant :

$$W_j^i = \begin{cases} H(X_i, Y_j) & \text{si } |H(X_i, Y_j)| \leq g(i, j) \\ 0 & \text{si } |H(X_i, Y_j)| > g(i, j); \end{cases}$$

Si $q(n) = \text{Prob}\{|Z| > n\}$, on sait que $q(n)$ est le terme général d'une série convergente, mais

$$q[g(i, j)] = \text{Prob}\{W_j^i \neq H(X_i, Y_j)\}$$

est alors le terme général d'une série double convergente d'après le lemme précédent; donc les suites des W_j^i et des $H(X_i, Y_j)$ sont équivalentes.

Posons

$$V_j^i = W_j^i + E\{W_j^i\} - E\{W_j^i|Y_j\} - E\{W_j^i|X_i\},$$

Ces variables aléatoires remplissent la condition C_2 avec

$$(\sigma_j^i)^2 = E\{(V_j^i)^2\} \leq E\{W_j^i\}^2$$

comme on l'a vu à la page 36.

Soit maintenant

$$Q(n) = \frac{1}{n^2} \int_0^n z^2 d[1 - q(n)]$$

on pourra appliquer le corollaire 1 bis à la suite des V_j^i si l'on montre que $Q[g(i, j)]$ est le terme général d'une série double convergente; comme on ne peut ici appliquer le premier lemme on établit directement :

Lemme - "Si $E(Z)$ existe, $Q[g(i, j)]$ est le terme général d'une série double convergente, quelles que soient les fonctions $n(p)$ et $m(p)$ satisfaisant à C' ".

Il suffit de montrer que

$$W(k) = [m(k) n(k) - m(k-1) n(k-1)] Q[m(k) n(k)]$$

est le terme général d'une série convergente. Soit alors

$$v(k) = \int_{k-1}^k z d[1 - q(n)]$$

on a

$$Q(n) = \frac{1}{n^2} \left[\sum_1^n \int_{k-1}^k z^2 d[1 - q(n)] \right] \leq \frac{1}{n^2} \sum_1^n k v(k);$$

donc $Q(n)$ est le terme général d'une série convergente.

Considérons la somme :

$$\sum_1^N W(k) \leq \sum_{\substack{v \leq m(k) n(k) \\ k \leq N}} v v(v) \frac{m(k) n(k) - m(k-1) n(k-1)}{m^2(k) n^2(k)};$$

étant donné que $f(k) = m(k) n(k)$ est une fonction croissante de k , tendant vers l'infini, et que $h(k) = \frac{1}{f(k)}$ est le terme général d'une série convergente, on peut remplacer la dernière expression par :

$$\sum_v^N v v(v) \left\{ \sum_{k_0 \leq k \leq N} \left[h(k) - \frac{h^2(k)}{h(k-1)} \right] \right\}$$

si $v = f(k_0)$. Il vient finalement :

$$\sum_1^N W(k) \leq \sum_v^N v v(v) \left[\frac{1}{v} - h(N) \right]$$

et le lemme en découle immédiatement.

Nous pouvons appliquer le corollaire 1 bis à la suite des V_j^i et prouver ainsi donc que :

$$\frac{1}{n(p) m(p)} \sum_{\substack{i \leq n(p) \\ j \leq m(p)}} V_j^i \rightarrow 0 \quad \text{p. s.} \quad (3)$$

quand $p \rightarrow \infty$. Mais en se reportant à la page 45 on voit que l'on peut écrire

$$V_j^i = W_j^i - E\{W_j^i\} - W_1(X_1) - W_2(Y_1)$$

avec : $E\{W_1(X_1)\} = E\{W_2(Y_1)\} = 0$

ainsi dans l'expression (3) les termes correspondant à $W_1(X_1)$ et $W_2(Y_1)$ disparaissent presque sûrement d'après la loi forte des grands nombres et on a donc montré que, lorsque $p \rightarrow +\infty$:

$$\frac{1}{n(p) m(p)} \sum_{\substack{i \leq n(p) \\ j \leq m(p)}} W_j^i \rightarrow \lim_{k \rightarrow \infty} \left[\frac{1}{n(k) m(k)} \sum_{\substack{i \leq n(k) \\ j \leq m(k)}} E\{W_j^i\} \right]$$

il est alors facile de voir que cette dernière expression est égale à $E(Z)$.

8 - Théorème II.

"Si $E\{|Z|^{3/2} \log|Z|\}$ existe, on a :

$$(P_n \times Q_m)^*(H) \longrightarrow E(Z) \quad \text{p. s.}$$

quand m et n tendent indépendamment vers l'infini".

La démonstration est assez semblable à celle du théorème précédent; on définit les W_j^i de la même façon avec :

$$g(i, j) = \sup(i, j)^2.$$

On a ici encore une suite équivalente puisque la fonction $g(i, j)$ correspond au cas particulier du théorème précédent, pour lequel :

$$m(p) = n(p) = p.$$

On passe de même des W_j^i aux V_j^i et pour leur appliquer le critère 2, il suffit donc d'établir le :

Lemme - "Si $E(|Z|^{3/2} \log|Z|)$ existe, alors :

$$\frac{1}{[\sup(i, j)]^2 \inf(i, j)} \int_0^{[\sup(i, j)]^2} z^2 d[1 - q(z)]$$

est le terme général d'une série double convergente".

Soit
$$v(k) = \int_{k-1}^k z^{3/2} \log(z) d[1 - q(z)]$$

on a à étudier :

$$\sum_{i>j} \frac{1}{i^2 j} \left[\sum_1^{i^2} \frac{\sqrt{k} \cdot v(k)}{\log k} \right] = \sum_{\substack{i>j \\ 1 \leq k \leq i^2}} \frac{\sqrt{k}}{i^2 j \log k} v(k)$$

Sommons sur j puis sur i et en majorant, il vient finalement que la somme de la série double est inférieure à celle de la série simple

$$v(k) \cdot \frac{1 + \log \sqrt{k}}{\log k}$$

donc a fortiori à celle de la série $v(k)$ de somme finie par hypothèse.

La fin du raisonnement est identique à celui des pages 45-46.

9 - Théorème III.

"Si $E(Z^2)$ existe, quelles que soient les fonctions $n(p)$ et $m(p)$ satisfaisant à C' , on a :

$$(P_n \times Q_m)^* (H) \longrightarrow E(Z) \quad \text{p. c. s.}$$

quand $p \longrightarrow \infty$."

Démonstration.

Posons :

$$E\{H(X, Y)\} = a$$

$$E\{H(X, Y) | X\} = H_1(X)$$

$$E\{H(X, Y) | Y\} = H_2(Y)$$

et

$$K(x, y) = H(x, y) + a - H_1(x) - H_2(y).$$

Or si $E(Z^2)$ existe, $E\{H_1(X)^2\}$ et $E\{H_2(Y)^2\}$ existent ; en effet :

$$E\{H_1(X)^2\} = E\{H(X, Y) H(X, Y')\} \leq E\{H(X, Y)^2\}.$$

Il s'ensuit que dans l'expression de :

$$\frac{1}{n(p) m(p)} \sum_{\substack{i \leq n(p) \\ j \leq m(p)}} K(X_i, Y_j)$$

$$\text{les termes } \frac{1}{n(p)} \sum_i H_1(X_i) \quad \text{et} \quad \frac{1}{m(p)} \sum_j H_2(Y_j)$$

tendent presque sûrement vers a , quand $p \longrightarrow + \infty$.

D'autre part, les $K(X_i, Y_j)$ sont quadratiquement indépendants et l'on a donc :

$$E \left\{ \left[\frac{1}{n(p) \ m(p)} \sum_{\substack{i \leq n(p) \\ j \leq m(p)}} K(X_i, Y_j) \right]^2 \right\} < \frac{E(Z^2)}{n(p) \ m(p)} ;$$

Comme sous la condition C', $\frac{1}{n(p) \ m(p)}$ est le terme général d'une série convergente, le théorème est établi.

10 - La convergence presque complètement sûre pour une suite double $V(m, n)$ sera entendue au sens suivant : la série double de terme général :

$$\text{Prob. } \{|V(m, n)| > \varepsilon\}$$

est convergente pour tout ε . Mais il découle immédiatement de cette définition que, pour tout n , $V(m, n)$ converge presque complètement sûrement vers 0 si $m \rightarrow \infty$; on ne peut donc prendre $(P_n \times Q_m)^* (H)$ pour établir un théorème de convergence presque sûre, mais on a seulement le théorème :

Théorème IV - "Si $E(|Z|^3)$ existe,

$$D(m, n) = (P_n \times Q_m)^* (H) - (P_x \times Q_m)^* (H) - (P_n \times P_y)^* (H) + E(Z)$$

tend vers 0, presque complètement sûrement, quand m, n tendent indépendamment vers l'infini".

Démonstration.

Il s'agit donc d'établir que

$$\text{Prob. } \{|D(m, n)| > \varepsilon\}$$

est le terme général d'une série double convergente.

On reprend à la fois les notations de la page 48 et celles de la page 36, et l'on a :

$$D(m, n) = \frac{1}{m \ n} \sum_{\substack{i \leq n \\ j \leq m}} K(X_i, Y_j) ;$$

Posons d'autre part

$$D'(m, n) = \frac{1}{m \ n} \sum_{\substack{i \leq n \\ j \leq m}} W(X_i, Y_j),$$

et montrons que

$$\text{Prob. } \{|D(m, n) - D'(m, n)| > \varepsilon\}$$

est le terme général d'une série double convergente; or il est facile de voir que cette probabilité est majorée par la somme des trois termes suivants :

$$mn \text{ Prob. } \{H_v(X, Y) \neq H(X, Y)\}$$

$$n \text{ Prob. } \{|V_1(X) - H_1(X) - a| > \varepsilon\}$$

$$m \text{ Prob. } \{|V_2(Y) - H_2(Y) - a| > \varepsilon\}.$$

Etant donné que $a - a_v \rightarrow 0$, quand m ou n tendent vers l'infini, on se ramène à étudier :

$$\text{Prob. } \{|E\{H_v(X, Y) - H(X, Y)|X\}| > \varepsilon\}$$

et l'expression analogue avec Y .

Utilisons l'inégalité de Markoff et remarquons que :

$$E\{|E\{H_v(X, Y) - H(X, Y)|X\}|\}$$

est inférieure à

$$E\{|H_v(X, Y) - H(X, Y)|\}$$

soit à

$$\int_{mn}^{\infty} z \, d[1 - q(z)].$$

Il nous reste donc à montrer, et nous le ferons à la fin de la démonstration pour des raisons de clarté, que :

$$mnq(mn) + (m + n) \int_{mn}^{\infty} z \, d[1 - q(z)]$$

est le terme général d'une série double convergente.

Nous établirons maintenant que $\text{Prob. } \{|D'(m, n)| > \varepsilon\}$ est le terme général d'une série double convergente en établissant qu'il en est de même pour :

$$E\{|D'(m, n)|^4\}.$$

Un calcul élémentaire, mais long montre que :

$$\begin{aligned} m^4 n^4 E\{D'(m, n)^4\} &= mn E\{W(X, Y)^4\} + \\ &+ \frac{3mn(m-1)(n-1)}{2} [E\{W(X, Y)^2\}]^2 \\ &+ 3mn(n-1) E\{[W(X, Y) W(X', Y)]^2\} \\ &+ 3mn(m-1) E\{[W(X, Y) W(X, Y')]^2\} \\ &+ mn(n-1)(m-1) E\{W(X, Y) W(X, Y') W(X', Y) W(X', Y')\}, \end{aligned}$$

où X' et X sont deux variables aléatoires indépendantes de loi P_X , et de même Y' et Y , deux variables aléatoires indépendantes de loi P_Y . En utilisant les majorations trouvées à la page 36, il vient aisément :

$$E\{D'(m, n)^4\} \leq \frac{6(m+n)-5}{m^3 n^3} E\{W(X, Y)^4\} + \frac{5(n-1)(m-1)}{2n^3 m^3} E\{H(X, Y)^2\},$$

le dernier terme fournit une série double convergente et nous avons seulement à étudier :

$$\frac{n+m}{n^3 m^3} E\{W(X, Y)^4\}.$$

Notons que :

$$E\{[V_1(X)]^4\} \quad \text{et} \quad E\{[V_2(Y)]^4\}$$

sont inférieurs à :

$$E\{[H_v(X, Y)]^4\},$$

et nous devons encore établir que

$$\frac{m+n}{n^3 m^3} \int_0^{mn} z^4 d[1 - q(z)]$$

est le terme général d'une série convergente, ce que nous ferons en même temps que les démonstrations exigées à la page 50.

Posons

$$u_k = \int_{k-1}^k z^3 d[1 - q(z)],$$

Comme $E\{|Z|^3\}$ existe, u_k est le terme général d'une série convergente et on a les majorations suivantes :

$$\int_{k-1}^k z^4 d[1 - q(z)] < k u_k$$

$$\int_k^{k+1} z d[1 - q(z)] < \frac{u_{k+1}}{k^2}$$

$$\int_k^{k+1} d[1 - q(z)] < \frac{u_{k+1}}{k^3}.$$

Il vient finalement :

$$\sum_{m,n} \frac{m+n}{m^3 n^3} \int_0^{mn} z^4 d[1 - q(z)] < \sum_{mn > k} k u_k \frac{m+n}{m^3 n^3}$$

$$\sum_{m,n} (m+n) \int_{mn}^{\infty} z d[1 - q(z)] < \sum_{mn < k} (m+n) \frac{u_{k+1}}{k^2}$$

et

$$\sum_{m,n} mn \int_{mn}^{\infty} d[1 - q(z)] < \sum_{mn < k} mn \frac{u_{k+1}}{k^3},$$

et donc, en intervertissant l'ordre des sommations, la démonstration est complète si on établit que :

$$\sum_{mn > k} \frac{m+n}{m^3 n^3} = O\left(\frac{1}{k}\right)$$

$$\sum_{mn < k} mn = O(k^3)$$

$$\sum_{mn < k} (m+n) = O(k^2)$$

et ceci est un calcul élémentaire.

B - VARIABLES ALÉATOIRES ENCHAÎNÉES, TYPE B

1 - Au paragraphe précédent on a pu remarquer que les conditions imposées pour la convergence de

$$(P_n \times Q_m)^* (H)$$

ne portaient que sur la loi de $Z = H(X, Y)$; or ici il n'en sera certainement pas de même : en effet, écrivons :

$$S_n = \frac{1}{n^2} \sum_{i \neq j \leq n} H(X_i, Y_j) + \frac{1}{n^2} \sum_{i=1}^n H(X_i, X_i).$$

Donc, S_n fait intervenir non seulement des variables aléatoires liées ayant pour loi celle de $Z' = H(X, X')$, mais aussi des variables aléatoires ayant pour loi celle de $Z_0 = H(X, X)$. Or il est évident que les conditions sur la loi de Z' et sur celle de Z_0 seront complètement indépendantes.

D'ailleurs ce phénomène n'a rien qui doive surprendre; en effet, dans l'un et l'autre de ces deux paragraphes, on a une estimation de la loi de Z ou Z' ; or dans le premier cas on dispose de deux échantillons indépendants pour estimer respectivement la loi de deux variables aléatoires indépendantes, tandis que dans le second cas, c'est le même échantillon qui sert à estimer la loi de deux variables aléatoires indépendantes; si d'ailleurs on disposait dans ce dernier cas de deux échantillons indépendants pour X , on retomberait dans le cadre de l'étude du précédent paragraphe; mais la suite montrera qu'il y a intérêt à les grouper en un seul échantillon.

On est donc conduit à séparer l'étude de

$$T_n = \frac{1}{n(n-1)} \sum_{i \neq j \leq n} H(X_i, X_j)$$

de celle de

$$R_n = \frac{1}{n^2} \sum_i H(X_i, X_i),$$

et il suffira de réunir les conditions de convergence obtenues séparément, pour avoir les conditions de convergence de

$$(P_n \times P_n)^* (H).$$

2 - L'étude de R_n , découle des théorèmes classiques, il ne peut y avoir convergence stochastique vers 0, que dans les conditions suivantes :

(A) $R_n \longrightarrow 0$ en probabilité, si et seulement si :

$$n [1 - K_0(n^2) + K_0(-n^2)] \longrightarrow 0$$

et

$$\frac{1}{n} \int_{-n^2}^{+n^2} z \, dK_0(z) \longrightarrow 0$$

quand $n \longrightarrow +\infty$, [8], où $K_0(z)$ est la loi de $H(X, X)$.

(B) $R_n \longrightarrow 0$ p.s. si et seulement si

$$E \{ |Z_0|^{\frac{1}{2}} \} < \infty.$$

(C) $R_n \longrightarrow 0$ p.c.s. si et seulement si

$$E \{ |Z_0| \} < \infty.$$

Pour cette dernière proposition il suffit de reprendre la démonstration donnée dans Dugué [2], pour la loi presque complètement sûre des grands nombres, en tronquant au niveau n^2 et en notant que si $q(n)$ est le terme décroissant d'une série positive convergente, $n q(n^2)$ est aussi le terme général d'une série convergente.

3 - Pour l'étude de T_n , on peut reprendre presque mot à mot les démonstrations du chapitre précédent; nous ne donnerons que celles qui ne s'adaptent pas directement.

Loi faible - "Si $n [1 - K'(n) + K'(-n)] \longrightarrow 0$

et
$$\int_{-n}^{+n} z \, dK'(z) \longrightarrow a$$

quand $n \longrightarrow +\infty$, T_n tend vers a en probabilité quand $n \longrightarrow +\infty$."

On se reporte à la page 36, $v = n^2$, la seule modification se trouve dans le calcul de

$$E \left\{ \sum_{i,j} W(X_i, X_j) \right\}^2 = n(n-1) E \{ W(X, X')^2 \} + n(n-1) E \{ W(X, X') W(X', X) \}$$

Mais ce dernier terme est inférieur à

$$n(n-1) E\{W(X, X')^2\}$$

et la suite est sans changement.

4 - Lois fortes.

On a ici des suites à un seul indice et il vient :

Lemme 1 - "Soit une suite $\{X_i\}$ de variables aléatoires indépendantes et de même loi; à tout couple (X_i, X_j) on fait correspondre fonctionnellement la variable aléatoire V_j^i telle que

$$(C_2^i) \quad E\{V_j^i\} = E\{V_j^i | X_i\} = E\{V_j^i | X_j\} = 0 \quad i \neq j$$

$$\text{et} \quad E\{V_i^i\} = 0.$$

$$\text{Soient} \quad S_n = \sum_{i,j \leq n} V_j^i$$

$$\text{et} \quad (\sigma_j^i)^2 = E\{(V_j^i)^2\},$$

on a :

$$\text{Prob.} \{ \sup |S_p| > \varepsilon, \quad p = 1, \dots, n \} \leq \frac{2}{\varepsilon^2} \sum_{i,j \leq n} (\sigma_j^i)^2 \quad ''.$$

Le raisonnement est le même qu'à la page 38 ; il faut cependant noter que les V_j^i sont quadratiquement indépendantes, sauf V_j^i et V_i^j , ce qui donne l'égalité :

$$E\{S_n^2\} = 2 \sum_{i,j \leq n} (\sigma_j^i)^2$$

car on sait que :

$$2 E\{V_j^i V_i^j\} \leq E\{(V_j^i)^2\} + E\{(V_i^j)^2\}.$$

5 - Critère 1.

"Si $\frac{(\sigma_j^i)^2}{[\text{Sup}(i,j)]^4}$ est le terme général d'une série double convergente, on a :

$$\frac{1}{n^2} S_n \rightarrow 0 \quad \text{p. s.}$$

quand $n \rightarrow +\infty$ ''.

On peut ici encore conduire la démonstration de deux façons différentes, celle de la page 39 se reproduit intégralement; pour l'autre il vient :

Critère 1 bis - Si $(\sigma_j^i)^2$ est le terme général d'une série double convergente, S_n admet presque sûrement une limite finie quand $n \rightarrow +\infty$.

6 - Théorème I.

"Si $E\{|Z'|\}$ existe,

$$T_n \rightarrow E\{Z'\} \quad \text{p. s.}$$

quand $n \rightarrow \infty$ ".

Reportons-nous à la page 44; prenons $g(i, j)$ égale à $\sup(i, j)^4$; les deux lemmes sont valables puisque cette fonction $g(i, j)$ correspond au cas particulier

$$n(p) = m(p) = p$$

et la démonstration est conservée exactement, mutatis mutandis, avec $V_j^i = 0$ pour $i \neq j$.

Théorème II. "Si $E\{Z'^2\} < \infty$,

$$T_n \rightarrow E\{Z'\} \quad \text{p. c. s.}$$

quand $n \rightarrow \infty$ ".

Les notations sont les mêmes qu'à la page 48, on remarque encore que

$$E\left\{\left[\sum_{i \neq j \leq n} K(X_i, X_j)\right]^2\right\} \leq 2n(n-1) E\{K(X, X')^2\}$$

et on conclut de la même manière.

C - EXTENSION AU CAS DE PLUSIEURS ECHANTILLONS

1 - Dans l'étude des produits du type

$$(P_{n_1} \times P_{n_2} \times \dots \times P_{n_p})^*$$

on se limitera pour simplifier l'exposé, aux deux cas suivants :

Cas A - On a p échantillons indépendants :

$$\{X_{i_k}^k\} \quad k = 1, \dots, p$$

$$i_k = 1, \dots, n_k$$

correspondant respectivement aux variables aléatoires X^k de lois P_k ; et comme précédemment on note P_{n_k} les diverses répartitions empiriques construites sur ces échantillons.

Cas B - On a 1 échantillon $\{X_i\}$ $i = 1, \dots, n$ de la variable aléatoire $\bar{X}$; l'étude faite au paragraphe précédent montre que lorsqu'on fait le produit $P_n \times P_n$, on doit se limiter à ce que nous appellerons le "produit restreint" et qui consiste à attribuer la masse $\frac{1}{n(n-1)}$ aux points (X_i, X_j) pour $i \neq j$ et la masse 0 aux points (X_i, X_i) . Plus généralement, le produit restreint d'ordre p pour P_n ($n > p$) sera la distribution aléatoire discrète dans R^p qui donne la masse $\frac{1}{n(n-1) \dots (n-p+1)}$ aux points $(X_{i_1}, \dots, X_{i_p})$ tels que les indices $i_1, \dots, i_p$ soient tous distincts; nous noterons cette opération $\times'$.

Dans ces deux cas, comme pour les paragraphes précédents, on a seulement à étudier la convergence des fonctionnelles associées aux répartitions empiriques. On peut démontrer les mêmes théorèmes que pour le cas $p = 2$ et nous nous bornerons à indiquer les principaux.

2 - Théorème 1.

"Si l'on a :

$$(C_1) \quad n[1 - K(n) + K(-n)] \longrightarrow 0$$

$$\text{et} \quad \int_{-n}^{+n} z \, dK(z) \longrightarrow a$$

quand $n \longrightarrow \infty$, alors

$$(P_{n_1} \times P_{n_2} \times \dots \times P_{n_p})^* (H) \longrightarrow a$$

en probabilité, quand $n_k \longrightarrow \infty$ pour tout k (cas A)

$$\text{ou} \quad (P_n \times' P_n \times' \dots \times' P_n)^* (H) \longrightarrow a$$

en probabilité quand $n \longrightarrow \infty$, (cas B) "

H est une fonction borélienne dans R^p et $K(z)$ désigne dans le cas A, la loi de la variable aléatoire

$$Z = H(X_1^1, \dots, X_p^p)$$

dans le cas B, la loi de la variable aléatoire

$$Z = H(X_1, \dots, X_p).$$

Démonstration.

$$\text{Soit } \nu = \begin{cases} n(n-1) \dots (n-p+1) & \text{dans le cas B} \\ \prod_{k=1}^p n_k & \text{dans le cas A} \end{cases}$$

dans les deux cas nous posons :

$$H_\nu(x_1, \dots, x_p) = \begin{cases} H(x_1, \dots, x_p) & \text{si } |H(x_1, \dots, x_p)| < \nu \\ 0 & \text{si } |H(x_1, \dots, x_p)| \geq \nu \end{cases}$$

$$a_\nu = \int_{-\nu}^{+\nu} z dK(z),$$

$$H_\nu(X_j) = E\{H_\nu(X^1, \dots, X^p) - a_\nu | X_j\}$$

(ou $H_\nu(X^j)$ dans le cas A), de même les espérances conditionnelles à plusieurs variables ; et enfin

$$W(X_1, \dots, X_p) = H_\nu(X_1, \dots, X_p) - a_\nu - \dots (-1)^{p-1} \sum_{j=1}^p H_\nu(X_j).$$

On note que dans les deux cas on a

$$E\{H_\nu(X_j)\} = 0$$

et que d'après (C_1) on peut se limiter aux fonctions H_ν ; il vient alors :

dans le cas (A) :

$$E \left\{ \left[\frac{1}{\nu} \sum_{i_1 \leq \dots \leq i_p} W(X_{i_1}^1, \dots, X_{i_p}^p) \right]^2 \right\} = \frac{1}{\nu} E \{ [W(X^1, \dots, X^p)]^2 \}$$

et l'on a :

$$E \{ [W(X^1, \dots, X^p)]^2 \} = \sigma^2\{H_\nu(X^1, \dots, X^p)\} - \dots \sum_{j=1}^{j=p} \sigma^2\{H_\nu(X_j)\}$$

ce qui est donc inférieur à

$$\int_{-\nu}^{+\nu} z^2 dK(z)$$

et on conclut comme à la page 36.

dans le cas (B) :

$$\begin{aligned} E \left\{ \left[\frac{1}{\nu} \sum_{i_k \leq n} W(X_{i_1}, \dots, X_{i_p}) \right]^2 \right\} &= \frac{1}{\nu} [E \{W(X_1, \dots, X_p)^2\} \\ &+ \sum_{j_1, \dots, j_p} E \{W(X_1, \dots, X_p) W(X_{j_1}, \dots, X_{j_p})\}] \end{aligned}$$

où $j_1, \dots, j_p$ est une permutation des nombres 1, 2, ..., p. On en déduit encore :

$$E \left\{ \left[\frac{1}{\nu} \sum_{i_k \leq n} W(X_{i_1}, \dots, X_{i_p}) \right]^2 \right\} \leq \frac{p!}{\nu} E \{W(X_1, \dots, X_p)^2\}$$

avec

$$E \{W(X_1, \dots, X_p)^2\} = \sigma^2 \{H_\nu(X_1, \dots, X_p)\} - \sum_{j=1}^p E \{H_\nu(X_j)^2\}$$

inférieur à

$$\int_{-\nu}^{+\nu} z^2 dK(z)$$

et on conclut comme précédemment.

3 - Pour les lois fortes donnons sans démonstration deux propositions qui généralisent le lemme 1.

Lemme 1_A - "Soient p suites de variables aléatoires indépendantes $\{X_{i_k}^k\}$; à tout ensemble $(X_{i_1}^1, \dots, X_{i_p}^p)$ on fait correspondre fonctionnellement la variable aléatoire $V(i_1, \dots, i_p)$ de telle façon que

$$E \{V(i_1, \dots, i_p) | X_{i_k}^k, \dots, X_{i_{k'}}^{k'}\} = 0 \quad \forall \quad k, k'$$

Supposons d'autre part que les n_k soient fonction d'un paramètre t, entier, de telle façon que :

$$(C') \quad \begin{cases} n_k \rightarrow \infty & \text{si } t \rightarrow \infty \\ n_k(t+1) > n_k(t) & \text{pour au moins un } k. \\ \text{la série } \frac{1}{\prod_k n_k(t)} \text{ est convergente.} \end{cases}$$

Alors si l'on pose :

$$S(t) = \sum_{i_k \leq n_k(t)} V(i_1, \dots, i_p)$$

on a

$$\text{Prob.} \{ \sup |S(t)| > \varepsilon, \quad t = 1, \dots, T \} \leq \frac{1}{\varepsilon^2} \sum_{i_k \leq n_k(t)} \sigma^2(i_1, \dots, i_p)$$

$$\text{avec} \quad \sigma^2(i_1, \dots, i_p) = E\{V(i_1, \dots, i_p)^2\} \quad ''.$$

Lemme 1₈ - "Soit une suite de variables aléatoires indépendantes $\{X_i\}$; à tout ensemble $(X_{i_1}, \dots, X_{i_p})$ on fait correspondre fonctionnellement la variable aléatoire $V(i_1, \dots, i_p)$ de telle façon que

$$E\{V(i_1, \dots, i_p) | X_{i_k}, \dots, X_{i_k}\} = 0$$

si tous les indices ne sont pas identiques et

$$E\{V(i, \dots, i)\} = 0 \quad \forall i.$$

Si

$$S_n = \sum_{i_k \leq n} V(i_1, \dots, i_p),$$

$$\text{on a} \quad \text{Prob.} \{ \sup |S_n| > \varepsilon, \quad n = 1, \dots, N \} \leq \frac{1}{\varepsilon^2} E\{S_n^2\} \quad ''.$$

Ces propositions sont ici encore à la base de la démonstration du théorème de convergence presque sûre :

4 - Théorème II.

"Si $E(|Z|)$ existe,

$$(P_{n_1} \times \dots \times P_{n_p})^* (H) \xrightarrow{\quad} E(Z) \quad \text{p. s.}$$

quand $t \rightarrow \infty$ (condition C')

$$\text{et} \quad (P_n \times \dots \times P_n)^* (H) \xrightarrow{\quad} E(Z) \quad \text{p. s.}$$

quand $n \rightarrow \infty$ ''.

Démonstration - Dans le cas (A), soit I_t l'ensemble d'indices

$$(i_1, \dots, i_p) \quad \text{tel que} \quad i_k \leq n_k(t)$$

$$\text{et soit} \quad K(i_1, \dots, i_p) = t \quad \text{si} \quad (i_1, \dots, i_p) \in I_t \cap I_{t-1}$$

et soit enfin

$$g(i_1, \dots, i_p) = \prod_{k=1}^p n_k [K(i_1, \dots, i_p)] .$$

On tronque au niveau $g(i_1, \dots, i_p)$ et la démonstration des pages 44 et suivantes se reproduit intégralement.

Dans le cas (B), on tronque au niveau $\sup(i_1, \dots, i_p)$ et il suffit de constater encore, que c'est un cas particulier de la fonction $K(i_1, \dots, i_p)$ précédente.

CHAPITRE V

THÉORÈME CENTRAL

LIMITE POUR DES VARIABLES ALÉATOIRES INTERDÉPENDANTES PLURI-DIMENSIONNELLES

On établit ici, et directement dans le cas de plusieurs échantillons, les propositions avancées au chapitre III, § C.

1 - Théorème I_A.

"Soient r fonctions boréliennes $H_1, \dots, H_j, \dots, H_r$ sur R^p et p échantillons $\{X_{i_k}^k\} : k = 1, \dots, p ; i_k = 1, \dots, n_k ;$

si :
$$E\{[H_j(X^1, \dots, X^p)]^2\} < \infty \quad \forall j$$

et si les $n_k(t)$ satisfont à la condition :

$$(C'') \quad \frac{\inf_k (n_k)}{n_k} \longrightarrow a_k \quad \text{quand} \quad t \longrightarrow \infty$$

alors le vecteur aléatoire $\vec{Z}(t)$ de R^r de composantes :

$$Z_j(t) = \frac{\sqrt{\inf_k (n_k)}}{\prod_k n_k(t)} \sum_{i_k \leq n_k} \{H_j(X_{i_1}^1, \dots, X_{i_p}^p) - E\{H_j(X_{i_1}^1, \dots, X_{i_p}^p)\}\}$$

tend en loi, quand $t \longrightarrow \infty$, vers le vecteur aléatoire laplacien de loi définie par :

$$E\{Z_j Z_{j'}\} = \sum_{k=1}^p a_k [E\{H_j(X^1, \dots, X^p) H_{j'}(X^1, \dots, X^p) - E\{H_j\} E\{H_{j'}\}}]''.$$

Démonstration :

Soit :

$$H_j^k(X^k) = E\{H_j(X^1, \dots, X^p) | X^k\} - E\{H_j(X^1, \dots, X^p)\}$$

et :

$$W_j(X^1, \dots, X^p) = H_j(X^1, \dots, X^p) - E\{H_j\} - \sum_{k=1}^p H_j^k(X^k).$$

Montrons que le vecteur de composantes :

$$W_j(t) = \frac{\sqrt{\inf_k (n_k)}}{\prod_k n_k} \sum_{i_k \leq n_k} W_j(X_{i_1}^1, \dots, X_{i_p}^p)$$

tend en probabilité vers 0 ; pour cela il suffit que :

$$E\{[W_j(t)]^2\} \longrightarrow 0 \quad \text{si } t \longrightarrow \infty.$$

Or :

$$E\{[W_j(t)]^2\} = \frac{\inf_k (n_k)}{\prod_k n_k^2} E\left\{ \left[\sum_{i_k \leq n_k} W_j(X_{i_1}^1, \dots, X_{i_p}^p) \right]^2 \right\}$$

soit encore :

$$\frac{\inf_k (n_k)}{\prod_k n_k} E\{[W(X^1, \dots, X^p)]^2\} + \sum_k 0 \left(\frac{1}{n_k} \right)$$

expression qui tend bien vers 0 si $t \longrightarrow \infty$.

La loi limite du vecteur $\vec{Z}(t)$ est donc identique à celle du vecteur de composantes :

$$Z_j^i(t) = \sum_{k=1}^p \left\{ \frac{\sqrt{\inf_k n_k}}{n_k} \sum_{i_k=1}^{n_k} H_j^k(X_{i_k}^k) \right\}$$

et le théorème I_A découle alors du théorème central limite classique.

Remarque :

$$E(Z_j^2) \leq \sum_{k=1}^{k=p} E\{H_j^k(X^k)^2\}$$

Car tous les a_k sont inférieurs ou égaux à 1. On voit que cette dernière somme est toujours inférieure à $\sigma^2(H_j)$; d'où :

$$\sigma^2(Z_j) \leq \sigma^2(H_j)$$

2 - On démontre de la même façon que :

Théorème I₈ -

"Soient r fonctions boréliennes $H_1, \dots, H_r$ sur R^p et un échantillon de dimension n de la variable aléatoire X .

Si :
$$E\{[H_j(X_1, \dots, X_p)]^2\} < \infty \quad \forall j,$$

le vecteur aléatoire de R^r de composantes :

$$Z_j(n) = \frac{\sqrt{n}}{n(n-1) \dots (n-p+1)} \sum_{j_k \text{ distincts, } \leq n} \{H_j(X_{i_1}, \dots, X_{i_p}) - E(H_j)\}$$

tend en loi quand $n \rightarrow \infty$, vers un vecteur aléatoire laplacien centré de loi définie par :

$$E\{Z_j Z_{j'}\} = E\{\hat{H}_j(X) \hat{H}_{j'}(X)\}$$

où :

$$\hat{H}_j(X) = \sum_1^p H_j^k(X)$$

et :

$$H_j^k(X) = E\{H_j(X_1, \dots, X_p) | X^k\}."$$

En effet on a avec les mêmes notations :

$$E\{[\sum W(X_{i_1}, \dots, X_{i_p})]^2\} \leq p! \cdot n(n-1) \dots (n-p+1) E\{W_j(X_1, \dots, X_p)^2\}$$

quantité qui tend vers 0 quand $n \rightarrow \infty$; on est ramené à la loi limite du vecteur de composantes :

$$\frac{\sqrt{n}}{n-p+1} \sum_{i=1}^{i=n} \hat{H}_j(X_i),$$

qui satisfait au théorème Central Limite.

FONCTION DETERMINANTE DE LA LOI D'ERREUR

1 - Soit d'abord une estimation de type A et d'ordre p de la loi $K(z)$; par la transformation de la page 25, qui ne dépend que de $K(z)$, on se ramène au cas, auquel nous nous limiterons dans tout ce paragraphe, où $K(z)$ est uniforme sur $(0,1)$. On étudie donc la loi limite de la fonction aléatoire :

$$\sqrt{\inf_k (n_k(t))} \quad [K_t(z) - z]$$

quand $t \rightarrow +\infty$.

Mettons en évidence une fonction, qu'on appellera le noyau $K(u, v)$, et qui déterminera tous les résultats ultérieurs. Soit $F_k(u, v)$ la loi de probabilité du couple des variables aléatoires :

$$U = H(X^1, \dots, X^p),$$

et :

$$V = H(X_1^1, \dots, X_1^{k-1}, X^k, X_1^{k+1}, \dots, X_1^p);$$

cette fonction de répartition $F_k(u, v)$ est symétrique et satisfait à :

$$F_k(u, 1) = u \quad F_k(1, v) = v.$$

la fonction :

$$F(u, v) = \frac{1}{\sum_k a_k} \sum_k a_k F_k(u, v)$$

possède les mêmes propriétés et peut être considérée comme une loi de probabilité dans le carré $0 \leq u, v \leq 1$.

2 - Le noyau :

$$K(u, v) = F(u, v) - uv$$

engendre une transformation symétrique et complètement continue A de $L^2_{(0,1)}$; si l'on note A_0 la transformation engendrée par le noyau :

$$K_0(u, v) = \inf(u, v) - uv,$$

A satisfait à la relation :

$$0 < a(Af, f) \leq (A_0 f, f) \quad (1)$$

où l'on a posé :

$$a = \sum_k a_k.$$

On peut écrire :

$$(Af, f) = \iint_{0 \leq u, v \leq 1} f(u) f(v) K(u, v) du dv$$

sous la forme :

$$\iint_{0 \leq u, v \leq 1} \left[\int_u^1 f(t) dt \right] \left[\int_v^1 f(t) dt \right] dK(u, v) ;$$

soient alors :

$$H_f(x_1, \dots, x_p) = \int_{H(x_1, \dots, x_p)}^1 f(t) dt - \int_0^1 t f(t) dt,$$

et :

$$H_f^k(X^k) = E\{H_f(X^1, \dots, X^p) | X^k\};$$

il vient :

$$a(Af, f) = \sum_{k=1}^p a_k E\{[H_f^k(X^k)]^2\}$$

et (Af, f) est strictement positif pour $f \neq 0$ p.p ; donc $K(u, v)$ est une covariance. D'autre part, la remarque de la page 64 entraîne l'autre partie de l'inégalité (1) en écrivant :

$$(A_0 f, f) = \int_0^1 \left[\int_u^1 f(t) dt \right]^2 du - \left[\int_0^1 t f(t) dt \right]^2.$$

Enfin $F(u, v)$ étant une loi de probabilité à lois marginales uniformes on a :

$$|F(u+h, v+k) - F(u, v)| < h+k,$$

donc $K(u, v)$ est continu et on peut appliquer le théorème de Mercer ; il existe donc un système orthonormal complet de $L_{(0,1)}^2$ tel que :

$$K(u, v) = \sum_{i=1}^{\infty} \mu_i \vartheta_i(u) \vartheta_i(v) ;$$

les ϑ_i sont continues, et si les μ_i sont indicées en ordre décroissants on a :

$$a\mu_i \leq \frac{1}{\pi^2 i^2}$$

3 - Soit maintenant le cas d'une estimation de type B et d'ordre p ; on met en évidence d'une façon analogue le noyau $K(u, v)$:

Soit $F_{k',k}(u, v)$ la loi de probabilité du couple :

$$U = H(X'_1, \dots, X'_{k-1}, X, X_{k+1}, \dots, X_p)$$

$$V = H(X'_1, \dots, X'_{k-1}, X, X_{k+1}, \dots, X'_p),$$

$F_{k',k}(u, v)$ n'est pas symétrique en général, mais :

$$F(u, v) = \frac{1}{p^2} \sum_{k',k} F_{k',k}(u, v)$$

est symétrique et représente une loi de probabilité sur le carré $0 \leq u, v \leq 1$, à lois marginales encore uniformes. Considérons de même la transformation associée au noyau :

$$F(u, v) - uv ;$$

on a ici encore :

$$0 < p(Af, f) \leq (A_0 f, f),$$

car $(Af, f) = E\{\hat{H}_f(X)^2\}$. $K(u, v)$ a donc ici les mêmes propriétés qu'à l'alinéa précédent, et sera dans la suite utilisé sans distinction dans le cas (A) ou (B).

4 - Le théorème, analogue au théorème de Prohorov pour des variables indépendantes, s'énonce ainsi :

Théorème - "La fonction déterminante de la loi d'erreur est une fonction aléatoire laplacienne, centrée, de covariance $K(u, v)$, presque sûrement continue et presque sûrement représentée dans $L^2_{(0,1)}$ par :

$$\Delta(t) = \sqrt{a} \sum_1^{\infty} \sqrt{\mu_1} U_1 \vartheta_1(t),$$

où : U_i sont des variables aléatoires normales centrées réduites et indépendantes,

où, $\{\vartheta_i\}$ est une base continue de $L^2_{(0,1)}$;

On a de plus :

$$a = \sum_k a_k \quad (\text{cas A})$$

$$a = p \quad (\text{cas B})$$

et si les μ_i sont rangés en ordre décroissant :

$$0 \leq \mu_i \leq \frac{1}{\pi^2 i^2} \quad (2)$$

La démonstration est analogue à celle donnée au chapitre II : on développe $\Phi_x(t) - t$ sur la base des ϑ_i et le théorème I_A ou I_B établit la proposition dans $L^2_{(0,1)}$. L'existence d'une fonction continue représentée dans $L^2_{(0,1)}$ par le développement trouvé découle du lemme utilisé dans la démonstration de Prohorov.

Remarque : L'inégalité (2) entraîne :

$$\forall x, \text{ Prob. } \left\{ \int_0^1 \Delta^2(t) dt > x \right\} < \text{Prob. } \{ w^2 > x \}$$

w^2 étant la variable aléatoire réduite de Von Mises. On notera qu'il n'est pas possible d'obtenir une inégalité analogue à la place du théorème de Kolmogorov-Smirnov.

4 - L'intérêt de cette méthode apparaît maintenant bien établi, par rapport à l'utilisation d'un échantillon de variables aléatoires indépendantes issues de l'ensemble étudié. Elle met de plus en évidence les covariances $K(u, v)$, où les lois de probabilité $F(u, v)$ sont symétriques, continues à lois marginales uniformes, sans que sans doute, ces propriétés soient caractéristiques de ces covariances.

On peut d'autre part, noter la cohérence des résultats précédents avec une remarque faite à la page 32 qui rappelle qu'ici, on peut écrire :

$$\Delta(t) = \sum_{k=1}^p a_k \Delta_k(t),$$

où les $\Delta_k(z)$ sont des fonctions aléatoires laplaciennes, indépendamment déduites des fonctions déterminantes de la loi d'erreur sur les variables aléatoires composantes. Cependant, l'application linéaire aléatoire associée à $\Delta(t)$ ne possède plus la forme (1) de la page 18, qui était indépendante de la base de $L^2_{0,1}$ choisie. On doit en effet considérer ici la forme quadratique :

$$(\alpha f, f) = \iint_{0 \leq u, v \leq 1} f(u) f(v) dK(u, v)$$

qui n'est pas strictement positive, mais qui satisfait, comme l'indique le calcul de la page 67, à :

$$0 \leq (\alpha f, f) \leq \int_0^1 f^2(x) dx - \left[\int_0^1 f(x) dx \right]^2.$$

le système $\{\vartheta_k^1\}$ de ses vecteurs propres n'est pas complet, mais joint à 1, il le devient, et l'on a pour $\Delta^*(\varphi)$ la représentation :

$$\Delta^*(\varphi) = \sum_{k=1}^{\infty} \sqrt{\mu_k^1} A_k < \varphi, \vartheta_k^1 >$$

où μ_k^1 est la valeur propre associée à ϑ_k^1 et $\{A_k\}$ un système de variables aléatoires normales centrées réduites et indépendantes.

5 - Application au test d'indépendance.

On peut songer à utiliser les résultats précédents pour construire un test d'indépendance sur le principe suivant :

Soient $\{(X_i, Y_i)\}$ $i = 1, \dots, n$, un échantillon dans le plan et $\varphi(x, y)$ une fonction borélienne arbitraire pour l'instant ; on a établi que, lorsque $n \rightarrow \infty$,

$$\sqrt{n} \left[\frac{1}{n^2} \sum_{i,j} \varphi(X_i, X_j) - E\{\varphi(X, Y)\} \right]$$

tend en loi vers une loi normale si et seulement si :

$$E\{\varphi^2(X, Y)\} < \infty.$$

Par conséquent la loi limite, quand $n \rightarrow \infty$, de :

$$T_n = \frac{\frac{1}{\sqrt{n}} \sum_{i,j} \varphi(X_i, Y_i) - \frac{1}{n\sqrt{n}} \sum_{i,j} \varphi(X_i, Y_j)}{\frac{1}{n} \sum_i \varphi^2(X_i, Y_i)}$$

est une loi normale : $N(0, \sigma)$ avec $\sigma \leq 1$, et ceci sous les conditions suivantes : X_i et Y_i indépendantes, $\varphi^2 \in C_{P_X \times P_Y}^2$.

On pourra donc reconnaître l'indépendance entre les deux

échantillons au fait que pour toutes les fonctions d'une classe Φ , la variable aléatoire T_n satisfait à ce comportement limite.

Il faut néanmoins reconnaître que cette classe Φ , aura sans doute la puissance du continu et rend ainsi l'investigation difficile, sans que l'indication fournie par quelques fonctions bornées soit pour autant négligeable.

CHAPITRE VI

CLASSES D'INVARIANCE DES FONCTIONS DÉTERMINANTES D'ERREUR DANS R^k .

L'étude d'échantillons, ou de "produits d'échantillons", a fait l'objet des chapitres précédents pour des variables aléatoires. Si l'on considère les mêmes problèmes pour des vecteurs aléatoires dans R^k , il semble s'introduire des complications d'écriture sans que les résultats essentiels soient modifiés. Un point essentiel, et d'ailleurs de grande importance, ne se transpose pas directement : c'est la forme réduite de la fonction déterminante d'erreur. Le changement de variable

$$y = F_x(x)$$

n'a plus de sens, ici. Plus précisément, bien qu'au chapitre II, la fonction aléatoire laplacienne qui avait la covariance

$$F_x[\inf(s, t)] = F_x(s) F_x(t)$$

et qui déterminait l'erreur statistique, soit simplement définie à partir de la fonction aléatoire de Wiener-Lévy, il s'est avéré utile, tant pour une démonstration en moyenne du théorème de Prohorov que pour celle du théorème de Von-Mises, de représenter cette fonction par ses composantes aléatoires sur la base de $L^2_{F_x}$ associée à sa covariance; c'est ce problème, élémentaire dans le cas de la droite que nous étudierons ici dans R^k .

1 - Soit donc P_z , une loi de probabilité sur R^k , et

$$F(x_1, \dots, x_k)$$

sa fonction de répartition; on étudie la fonction aléatoire laplacienne à k variables, $\Delta(x_1, \dots, x_k)$, de covariance :

$$\Gamma(x_1, \dots, x_k; y_1, \dots, y_k) = F(u_1, \dots, u_k) - F(x_1, \dots, x_k) F(y_1, \dots, y_k)$$

$$\text{où} \quad u_i = \inf(x_i, y_i) \quad i = 1, \dots, k$$

On désigne encore par $L_{P_z}^2$, l'espace de Hilbert des classes de fonctions boréliennes de carré sommable par rapport à P_z :

$$\int f^2 dP_z < \infty .$$

Dans toute la suite nous utiliserons les notations suivantes : si P et Q appartiennent à R^k , on écrira :

$$P < Q$$

si toutes les coordonnées de P , sont respectivement inférieures à celles de Q ; on écrira aussi

$$P \wedge Q$$

pour $\inf(P, Q)$ au sens de la relation d'ordre précédente.

Enfin χ désignera la fonction caractéristique du cône positif et ε , le point dont toutes les coordonnées sont $+\infty$, c'est-à-dire tel que

$$P \wedge \varepsilon = P \quad \forall P.$$

On peut donc écrire la covariance Γ , de la fonction aléatoire $\Delta(P) = \Delta(x_1, \dots, x_k)$ sous la forme :

$$\Gamma(P, Q) = F(P \wedge Q) - F(P) F(Q),$$

et on remarque, ceci découlant de la définition même de Δ , que la fonction aléatoire

$$\delta(P) = \chi(P - P^*)$$

a aussi Γ pour covariance, P^* désignant un point aléatoire de R^k de loi P_z . $\delta(P)$ appartenant à $L_{P_z}^2$, on a les propriétés classiques des fonctions aléatoires du second ordre :

(A) $\Gamma(P, Q)$ est un noyau complètement symétrique, défini positif, qui engendre dans $L_{P_z}^2$, la forme quadratique :

$$\langle \Gamma f, f \rangle = \int \Gamma(P, Q) f(P) f(Q) dF(P) dF(Q)$$

égale d'ailleurs à

$$\sigma^2(\langle f, \delta \rangle).$$

Les éléments propres, (s_n, φ_n) , de ces noyaux satisfont aux relations

$$v_n = \int \Gamma^{(n)}(P, P) dF(P) \quad n \geq 1 \quad (1)$$

$$\nu_n = \sum_i (s_i)^n,$$

$\Gamma^n(P, P)$ étant la trace du noyau itéré d'ordre n de Γ .

(B) La fonction aléatoire laplacienne $\Delta(P)$ admet, dans $L^2_{P_X}$, la représentation presque sûre :

$$\Delta(P) = \sum_{i=0}^{\infty} \sqrt{s_i} U_i \varphi_i(P) \quad (2)$$

où les U_i sont des variables aléatoires normales centrées réduites et indépendantes.

2 - On est donc amené pour expliciter cette représentation, à résoudre l'équation intégrale :

$$(E) \quad s f(P) = \int \Gamma(P, Q) f(Q) dF(Q)$$

Cette résolution suppose le calcul de la fonction de Fredholm $D(\lambda)$ du noyau Γ et il se trouve qu'ici cette fonction a de plus une interprétation probabiliste très importante :

Soient en effet $G(z)$ la fonction, entière d'après (1),

$$\prod_0^{\infty} \left(1 - \frac{z}{\lambda_i}\right) \quad \left(s = \frac{1}{\lambda}\right)$$

et

$$W^2 = \int \Delta^2(P) dF(P)$$

la variable aléatoire qui généralise le W^2 de Von-Mises. W est la norme de Δ dans $L^2_{P_X}$ et admet d'après (2) la représentation :

$$W^2 = \sum_0^{\infty} s_i U_i^2;$$

la fonction caractéristique de W^2 :

$$\varphi_{W^2}(t) = \prod_0^{\infty} \frac{1}{\sqrt{1 - 2it s_j}}$$

sera donc aussi égale à :

$$\frac{1}{\sqrt{D(2it)}}$$

puisque ici $D(\lambda)$ est de genre 0. Comme pour $|\lambda| < 1$, on a

$$\log D(\lambda) = - \sum_1^{\infty} \frac{\lambda^n}{n} v_n,$$

on passe aisément des cumulants de la distribution de W^2 aux traces des noyaux itérés de Γ , sans que pour autant ceux-ci soient explicitement calculables.

3 - L'étude qui nous occupera dorénavant seule dans toute la suite, de l'équation (E) sera simplifiée si l'on ramène (E) à une forme non singulière : écrivons pour cela E sous la forme :

$$s f(P) = \int F(P \wedge Q) f(Q) dF(Q) - F(P) \int F(\varepsilon \wedge Q) f(Q) dF(Q)$$

et soit (E_1) l'équation :

$$(E_1) \quad s f(P) = \int F(P \wedge Q) f(Q) dF(Q) + F(P),$$

E_1 a une solution unique $f_1(s, P)$ et on a la proposition :

Proposition - "Les valeurs propres du noyau Γ sont les racines non nulles de l'équation

$$f_1(s, \varepsilon) = 0; \quad (3)$$

si s_i est une racine de (3), la fonction propre correspondante est :

$$\varphi_i(P) = - f_1(s_i, P) \times \frac{1}{\sqrt{\int f_1^2(s, P) dF(P)}}."$$

En effet, reportons cette expression de φ_i dans (E); il vient au second membre :

$$- \frac{1}{||f_1||} \int F(P, Q) f_1(s_i, Q) dF(Q) + \frac{F(P)}{||f_1||} \int F(\varepsilon \wedge Q) f_1(s_i, Q) dF(Q) \quad (4)$$

où $f_1(s_i, Q)$ satisfait à (E_1) ; donc

$$s_i f_1(s_i, P) = \int f_1(s_i, Q) F(P \wedge Q) dF(Q) + F(P)$$

et en faisant $P = \varepsilon$, on a

$$\int f_1(s_i, Q) F(Q) dF(Q) = -1, \quad f_1(s_i, \varepsilon) \text{ étant nul.}$$

Soit en reportant dans (4), on a bien pour ce second membre

$$\frac{s_1 f_1(s_1, P) + F(P) - F(P)}{||f_1||} \quad \text{soit } \varphi_1(P)$$

Réciproquement

On passe sans difficulté de (E) à (E₁) sous la seule condition :

$$(E_2) \quad sf(P) = \int f(Q) F(P \wedge Q) dF(Q)$$

n'a pas de solutions de carré sommable, nulles pour $P = \varepsilon$.

Nous n'avons pas établi que cette condition était vérifiée pour toute loi de probabilité F, mais on peut facilement voir qu'elle est remplie pour les lois dont le support est totalelement ordonné par la relation d'ordre considérée. Notons d'autre part qu'en écrivant (E₂) sous la forme

$$sf(P) = \int_{R < P} \left\{ \int_{Q > R} f(Q) dF(Q) \right\} dF(R)$$

on peut se ramener à une équation aux dérivées partielles. Enfin, on se reportera à cette condition quand on aura introduit sur les lois F une relation d'équivalence précisée à la fin de ce chapitre.

A - FONCTIONS DE FREDHOLM D(λ) ET Δ(λ) DES NOYAUX Γ ET Γ₁ = F(P ∧ Q)

$$D(\lambda) = \sum_0^{\infty} \frac{(-\lambda)^p}{p!} \int |F(P_i \wedge P_j) - F(P_i) F(P_j)| dF(P_1) \dots dF(P_p)$$

$$i, j = 1 \dots p \quad (I-1)$$

$$\Delta(\lambda) = \sum_0^{\infty} \frac{(-\lambda)^p}{p!} \int |F(P_i \wedge P_j)| dF(P_1) \dots dF(P_p)$$

Bordons le déterminant $|F(P_i \wedge P_j) - F(P_i) F(P_j)|$ par une ligne supplémentaire $[F(P_1), \dots, F(P_p), 1]$ et une colonne supplémentaire $[0, \dots, 0, 1]$; en ajoutant la ligne supplémentaire précédente multipliée par $F(P_j)$ à chaque ligne j on obtient le nouveau déterminant $|F(P_i \wedge P_j)|$ $i, j = 1, \dots, p+1$ en posant $P_{p+1} = \varepsilon$.

Pour ce déterminant, comme pour celui de $\Delta(\lambda)$, on introduit pour chaque ligne i une variable Q_i de façon à écrire :

pour i fixé, $j = 1, \dots, p$ $F(P_i \wedge P_j) = \int \chi(P_i - Q_i) \chi(P_j - Q_i) dF(Q_i)$;

on peut alors intervertir le signe somme et le signe déterminant pour aboutir à la nouvelle forme :

$$D(\lambda) = \sum_0^{\infty} \frac{(-\lambda)^p}{p!} \int \left| \chi(P_i - Q_i) \chi(P_j - Q_i) \right| dF(P_1) \dots dF(P_p) dF(Q_1) \dots dF(Q_{p+1})$$

$$i, j = 1, \dots, p+1 \quad P_{p+1} = \varepsilon \quad (I-2)$$

$$\Delta(\lambda) = \sum_0^{\infty} \frac{(-\lambda)^p}{p!} \int \left| \chi(P_i - Q_i) \chi(P_j - Q_i) \right| dF(P_1) \dots dF(P_p) dF(Q_1) \dots dF(Q_p)$$

$$i, j = 1, \dots, p$$

Pour chacun des nouveaux déterminants on peut mettre dans chaque ligne $\chi(P_i - Q_i)$ en facteur pour $i = 1, \dots, p$. Compte tenu de ce que l'on a dans le domaine d'intégration $\chi(P_i - Q_i) = 1 \forall i$, il reste deux déterminants qui peuvent s'écrire : le premier : $|\chi(P_j - Q_i)|$ sauf sur la diagonale principale et la dernière colonne où les termes sont égaux à 1, le second étant obtenu en supprimant dans le premier la dernière ligne et la dernière colonne.

Dans ces deux déterminants on pose $\chi^* = 1 - \chi$; dans le premier on retranche la dernière colonne à toutes les autres et il vient $(-1)^p \delta_1$; dans le second on borde par une colonne $(1, \dots, 1)$ et une ligne $(0, \dots, 0, 1)$, on retranche cette dernière colonne à toutes les autres et il vient $(-1)^p \delta_2$. Les déterminants δ_1 et δ_2 sont obtenus à partir du déterminant

$$\delta_3 = \left| \chi^*(P_i - Q_j) \text{ pour } i \neq j, 0 \text{ pour } i = j; i, j = 1, \dots, p \right|$$

pour δ_1 , en bordant par la colonne $(1, \dots, 1)$ et la ligne $(\chi^*(P_1 - Q_{p+1}), \dots, \chi^*(P_p - Q_{p+1}), 1)$;

pour δ_2 , en bordant par une ligne et une colonne supplémentaire formées de 1.

On étudiera finalement le système :

$$D(\lambda) = \sum_0^{\infty} \frac{\lambda^p}{p!} A_p, \quad \Delta(\lambda) = \sum_0^{\infty} \frac{\lambda^p}{p!} B_p, \quad C(\lambda) = \sum_0^{\infty} \frac{\lambda^p}{p!} C_p$$

$$A_p = \int_{E_p} \delta_1(P_1, \dots, P_p; Q_1, \dots, Q_{p+1}) dF(P_1), \dots, dF(P_p) dF(Q_1), \dots, dF(Q_{p+1})$$

$$B_p = \int_{E_p} \delta_2(P_1, \dots, P_p; Q_1, \dots, Q_p) dF(P_1), \dots, dF(P_p) dF(Q_1), \dots, dF(Q_p) \quad (1.3)$$

$$C_p = \int_{E_p} \delta_3(P_1, \dots, P_p; Q_1, \dots, Q_p) dF(P_1), \dots, dF(P_p) dF(Q_1), \dots, dF(Q_p)$$

E_p est défini par

$$\prod_{i=1}^p \chi(P_i - Q_i) = 1.$$

Les fonctions $D(\lambda)$, $\Delta(\lambda)$, $C(\lambda)$ sont entières et nous nous proposons de les mettre sous une forme plus simple en établissant sur les suites des A_p , B_p , C_p , des relations de récurrence toutes trois obtenues en appliquant aux déterminants δ_1 , δ_2 , δ_3 , la formule de développement symétrique :

$$|a_j^i| = \sum_{k=1}^{k=n} (-1)^{k+1} \sum_{i_1, \dots, i_{k-1}} a_{i_1}^1 a_{i_2}^2, \dots, a_{i_{k-1}}^{k-1} \Delta_{1, i_1, \dots, i_{k-1}}$$

où $\Delta_{i, j, \dots, l}$ est le déterminant obtenu à partir de $|a_j^i|$ en supprimant les lignes et les colonnes d'indice $i, j, \dots, l$, et où la suite $i_1, \dots, i_{k-1}$ est une suite de $k-1$ indices distincts de $2, \dots, n$. De plus $\Delta_{1, i_1, \dots, i_{k-1}}$ sera noté 1. (On sait que cette formule est obtenue en classant les permutations de $1, \dots, n$ par la longueur du cycle contenant 1).

Relations de récurrence.

Quand dans δ_3 on supprime un certain nombre de lignes et de colonnes de même indice, on a encore un déterminant de la forme δ_3 soit

$$\delta_3(P_1, \dots, P_p; Q_1, \dots, Q_p) = \sum_{k=2}^{k=p} (-1)^{k+1} \sum_{i_1, \dots, i_{k-1}} \chi^*(P_1 - Q_{i_1}) \chi^*(P_{i_1} - Q_{i_2}), \dots, \chi^*(P_{i_{k-1}} - Q_1) \delta_3(P_i, Q_j; i, j \neq i_1, \dots, i_{k-1})$$

En intégrant sur E_p , il y a séparation des variables et il vient :

$$C_p = \sum_{k=2}^{k=p} (-1)^{k+1} C_{p-k} \frac{(p-1)!}{(p-k)!} \gamma_k \quad (C_0 = 1 \quad C_1 = 0)$$

où

$$\gamma_k = \int_{\chi(P_i - Q_i)=1, \chi(P_i - Q_{i+1})=0, i=1, \dots, k, Q_{k+1}=Q_1} dF(P_1) \dots dF(P_k) dF(Q_1) \dots dF(Q_k)$$

On peut procéder de la même façon pour δ_1 et δ_2 , mais on aura à tenir compte de ce que $p + 1$ appartient ou non à l'ensemble $(i_1, \dots, i_{k-1})$. Il vient

$$\delta_1(P_1, \dots, P_p; Q_1, \dots, Q_p) =$$

$$\sum_{k=2}^{k=p} (-1)^{k+1} \left\{ \sum_{\substack{i_1 \dots i_{k-1} \in 1 \dots p \\ i_j = p+1}} \chi^*(P_1 - Q_{i_1}) \dots \chi^*(P_{j-1} - Q_{i_j}) \chi^*(P_{j+1} - Q_{i_{j+2}}) \dots \right. \\ \left. \dots \chi^*(P_{i_{k-1}} - Q_{i_1}) \delta_1(P_1, Q_j, i, j \neq i_1 \dots i_{k-1}) + \right. \\ \left. + \sum_{i_1 \dots i_{k-1} \in 2 \dots p} \chi^*(P_1 - Q_{i_1}) \dots \chi^*(P_{i_{k-1}} - Q_{i_1}) \delta_3(P_1; Q_j, i, j \neq i_1 \dots i_{k-1}) \right\}$$

En intégrant il vient

$$A_p = \sum_{k=2}^{k=p+1} (-1)^{k+1} \left[A_{p-k} \frac{(p-1)!}{(p-k)!} \gamma_k + (k-1) \frac{(p-1)!}{(p-k+1)!} \alpha_k C_{p-k+1} \right]$$

$$A_0 = 1 \quad A_{-1} = 0 \quad B_0 = 1 \quad B_{-1} = 0$$

et de même

$$B_p = \sum_{k=2}^{k=p+1} (-1)^{k+1} \left[B_{p-k} \frac{(p-1)!}{(p-k)!} \gamma_k + (k-1) \frac{(p-1)!}{(p-k+1)!} \beta_k C_{p-k+1} \right]$$

avec

$$\alpha_k = \int_{\substack{\chi(P_i - Q_i) = 1, \quad i=1 \dots k-1 \\ \chi(P_{i-1} - Q_i) = 0, \quad i=2 \dots k}} dF(P_1) \dots dF(P_{k-1}) dF(Q_1) \dots dF(Q_k)$$

$$\beta_k = \int_{\substack{\chi(P_i - Q_i) = 1, \quad \chi(P_{i-1} - Q_i) = 0 \\ i=1 \dots k-1 \quad i=2 \dots k-1}} dF(P_1) \dots dF(P_{k-1}) dF(Q_1) \dots dF(Q_{k-1})$$

Chaînes et cycles de R^k -

Les suites numériques $\alpha_k, \beta_k, \gamma_k$ s'interprètent aisément si on introduit la notation de :

$$\underline{\text{cycle}} \quad \text{si} \quad P_1 \not\prec Q_k < P_k \not\prec Q_{k-1}, \dots, P_2 \not\prec Q_1 < P_1$$

chaîne impaire si $Q_1 < P_1 \nless P_2 < P_2, \dots, P_{k-1} \nless Q_k$

chaîne paire si $Q_1 < P_1 \nless Q_2 < P_2, \dots, P_{k-2} \nless Q_{k-1} < P_{k-1}$

En effet, $\alpha_k, \beta_k, \gamma_k$ sont alors respectivement les probabilités pour que des points aléatoires indépendants de loi F soient en cycle, chaîne paire ou impaire.

Les fonctions génératrices de ces probabilités :

$$\gamma(\lambda) = \sum_2^{\infty} (-\lambda)^{k-1} \gamma_k \quad \beta(\lambda) = \sum_2^{\infty} (-\lambda)^{k-1} \beta_k \quad \alpha(\lambda) = \sum_2^{\infty} (-\lambda)^{k-1} \alpha_k$$

ne sont pas entières mais ont un rayon de convergence au moins égal à 1 et on arrive par une transformation simple des équations de récurrence à :

$$|\lambda| < 1 \quad \left\{ \begin{array}{l} C'(\lambda) = \gamma(\lambda) C(\lambda) \\ D'(\lambda) = D(\lambda) \gamma(\lambda) + C(\lambda) \alpha'(\lambda) \\ \Delta'(\lambda) = \Delta(\lambda) \gamma(\lambda) + C(\lambda) \beta'(\lambda) \end{array} \right.$$

$C(\lambda), \Delta(\lambda)$ et $D(\lambda)$ étant entières, ceci donne pour tout λ

$$\begin{array}{l} D(\lambda) = [1 + \alpha(\lambda)] \exp. \int_0^{\lambda} \gamma(u) du \\ \Delta(\lambda) = [1 + \beta(\lambda)] \exp. \int_0^{\lambda} \gamma(u) du \end{array}$$

B - SOLUTION DE L'EQUATION E_1

Soient les fonctions :

$$G_n(P) = \text{Prob.} \{ Q_1 < P_1 \nless Q_2, \dots, Q_{n-1} < P \}$$

les Q_i et P_i étant des points aléatoires indépendants de loi F; soit

$$G(P) = \sum_2^{\infty} (-\lambda)^{n-2} G_n(P) \quad |\lambda| < 1 \quad (\text{II-1})$$

Théorème.

" $G(P)$ est une fonction méromorphe entièrement définie par (II-1) et la solution $f_1(\frac{1}{\lambda}, P)$ de E_1 s'écrit :

$$f_1\left(\frac{1}{\lambda}, P\right) = G(P) / \int G(P) dF(P) \quad \text{ou encore} \quad G(P) \exp. \int_0^{\lambda} \gamma(u) du / \Delta(\lambda)$$

On notera que le numérateur de cette dernière expression est entier".

Remarque - $f_1\left(\frac{1}{\lambda}, \varepsilon\right) = \frac{D(\lambda)}{\Delta(\lambda)}$ mais E_2 n'ayant pas de solutions, $D(\lambda)$ et $\Delta(\lambda)$ n'ont pas de racines communes.

Démonstration - D'après l'expression de $\Delta(\lambda)$, il suffit de montrer que $G(P)$ est solution de l'équation :

$$G(\lambda, P) = \lambda \int F(P \wedge Q) G(\lambda, Q) dF(Q) + F(P) [1 + \gamma(\lambda)]$$

Désignons par μ la mesure sur R^{2n-1} :

$$F(Q_1) \times \dots \times F(Q_n) \times F(P_1) \times \dots \times F(P_{n-1}),$$

on peut écrire le terme en $(-\lambda)^{n-1}$ du second membre comme suit :

$$\mu \{ Q_1 < P_1 \nmid Q_2, \dots, Q_{n-1} < P_{n-1}; Q_n < P \}$$

$$- \mu \{ Q_1 < P_1 \nmid Q_2, \dots, Q_{n-1} < P_{n-1}; Q_n < P_{n-1}, Q_n < P \},$$

On a pris seulement P_{n-1} au lieu de Q comme variable d'intégration. Les ensembles obtenus sont contenus le second dans le premier, leur différence étant définie par :

$$Q_1 < P_1 \nmid Q_2, \dots, Q_{n-1} < P_{n-1} \nmid Q_n < P;$$

on retrouve bien le terme en $(-\lambda)^{n-1}$ de $G(P)$, et comme de plus $G_2(P)$ est égal à $F(P)$ la proposition est établie.

C - CLASSES DE LOIS DE PROBABILITE SUR R^k

La relation d'ordre déterminant seule les solutions, soit T une transformation sur R^k qui conserve la relation d'ordre, c'est-à-dire telle que

$$\chi(P, Q) = \chi(TP, TQ),$$

et soit $\mathcal{T}$ la famille de ces transformations. Si $\mathcal{P}$ est l'espace des lois de probabilités sur R^k , introduisons sur $\mathcal{P}$ la relation d'équivalence

$$P_1 \sim P_2$$

s'il existe deux transformations T_1 et T_2 de $\mathfrak{C}$, telles que

$$F_1(M) = F_2(T_1 M) \quad \text{et} \quad F_2(M) = F_1(T_2 M)$$

$F_1(M)$ et $F_2(M)$ étant les fonctions de répartition de P_1 et P_2 ; si alors $\mathfrak{C}\mathfrak{M}$ est l'espace quotient formé des classes d'équivalence de $\mathfrak{M}$ on a le théorème suivant qui résume tout ce qui précède :

Théorème - "A chaque classe de $\mathfrak{C}\mathfrak{M}$ on peut faire correspondre:

1/ une fonction $G_\lambda(P)$ sur R^k , méromorphe en λ ;

2/ deux fonctions numériques $\gamma_0(\lambda)$ et $B(\lambda)$ telles que

$$B(\lambda), \quad \gamma_0(\lambda) B(\lambda) \quad \text{et} \quad G_\lambda(P) B(\lambda)$$

soient entières en λ . On a alors :

1/ les valeurs propres λ_i ne dépendent que de $\mathfrak{C}\mathfrak{M}$ et sont les racines des

$$G_\lambda(\varepsilon) = 0$$

2/ les fonctions propres $\varphi_i(P)$ correspondant à toutes les lois de probabilité d'une même classe de $\mathfrak{C}\mathfrak{M}$, sont obtenus quand T parcourt $\mathfrak{C}$ dans la formule

$$\varphi_i(P) = -\lambda_i \frac{G_{\lambda_i}(TP)}{\gamma_0(\lambda_i)} \quad ''.$$

Remarque.

1/ Dans le cas de la droite la relation d'ordre donnée par $\chi(P, Q)$ est totale; les lois de probabilité continues appartiennent donc à la même classe d'équivalence et les fonctions $B(\lambda)$, $\gamma_0(\lambda)$ et $G_\lambda(P)$ y sont uniques : il vient immédiatement

$$B(\lambda) = 1$$

$$\gamma_0(\lambda) = \cos \sqrt{\lambda}$$

$$G_\lambda(P) = \begin{cases} \frac{\sin x \sqrt{\lambda}}{\sqrt{\lambda}} & \text{pour } 0 \leq x \leq 1 \\ \text{constante} & \text{pour } x < 0 \text{ ou } x > 1. \end{cases}$$

On retrouve donc :

les valeurs propres $n^2\pi^2$ solutions de

$$\frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}} = 0$$

et les fonctions propres réduites :

$$(-1)^{n+1} \frac{\sin [T(x) \sqrt{\lambda_1}]}{\sqrt{\lambda_1}},$$

$T(x)$ est une fonction monotone de x et on voit que la fonction $F(x)$ de la théorie classique, doit être considérée moins comme la fonction de répartition de la loi de probabilité que comme une transformation.

2/ Si on se propose de calculer la distribution de

$$\int \Delta^2 d\mu$$

où μ est une loi de probabilité différente de P_z , la méthode précédente se généralise immédiatement : il suffira en effet de considérer que les points P_j suivent la loi P_z , les points Q_i la loi μ , tout en restant mutuellement indépendants. Ceci peut être particulièrement intéressant si une "fonction de prix" est par ailleurs introduite dans le problème.

D - GENERALISATION

Soit $(\mathcal{H}, \mathcal{A}, P)$ un espace vectoriel topologique ordonné $\mathcal{H}$, probabilité par la mesure P sur le σ -corps $\mathcal{A}$. Il doit exister entre $\mathcal{A}$ et les ouverts de la topologie sur $\mathcal{H}$ des rapports permettant :

1/ de parler de probabilité de "convergence" pour que l'on puisse songer à étendre les lois fortes des grands nombres;

2/ et surtout, ici, de définir P par sa fonction de répartition :

$$x \in \mathcal{H} \quad F_x(x) = P. \{y \in \mathcal{H} : y < x\},$$

le cône positif et négatif de chaque point doit donc être mesurable; on pourra trouver dans Revuz [7] les conditions nécessaires et suffisantes à adjoindre.

Soit alors P_n , une fonction de répartition empirique; l'étude de la loi limite de la fonction aléatoire sur $\mathcal{H}$

$$\sqrt{n} [F_n(x) - F_x(x)]$$

conduit aussi à résoudre dans L^2_p , l'équation intégrale :

$$f(x) = \lambda \int_{\mathcal{H}} F_x(x \wedge y) f(y) dP - \lambda F_x(x) \int_{\mathcal{H}} F_x(y) f(y) dP.$$

Sous les conditions de [7] on peut écrire dP ou $dF_x(x)$.

Cette équation se résoudra comme celle de la page 75 et le théorème fondamental se transpose directement.

Si donc l'étude d'une fonction de répartition empirique semble devoir se prolonger pour des E.V.T. ordonnés et probabilisés, satisfaisant à des conditions de compatibilité entre la topologie et la loi de probabilité, néanmoins, le problème qu'elle suggère des lois des grands nombres dans de tels espaces demeure entier.

BIBLIOGRAPHIE

Seules les publications numérotées sont citées dans le texte.

AGGARWAL - On some minimax invariant procedures for estimating a cumulative distribution function. - *Annals of Mathematical Statistics* (26), 1955, p.450.

1] BARRA - 1 - Du comportement de la moyenne d'un échantillon. - *C.R. Ac. Sc. Paris*, t.244, p.2002, (1957).

2 - Sur une propriété des fonctions de répartition empiriques considérées comme estimations des fonctions de répartition théoriques. - *C.r.Ac.Sc. Paris*, t.244, p.3020 (1957).

3 - Lois des grands nombres pour des suites doubles de variables aléatoires. - *C.R.Ac.Sc. Paris*, t.246, p.2999-3000 (1958).

4 - Extension du théorème de Von Mises à certaines classes de lois de probabilité sur R^k et plus généralement sur un espace de Riesz. - *C.R.Ac.Sc. Paris*, t.250, p.52, (1960).

2] DUGUE - 1 - Sur la convergence stochastique au sens de Césaro Sankha, 1957, Vol.18.

2 - Eléments limites stochastiques. - 28ième session de l'Institut International de Statistique.

DONSKER - Justification and extension of Doob heuristic approach to the Kolmogoroff-Smirnov theorem. - *Annals of Mathematical Statistics* 1952 (23) n° 2.

DOOB - Heuristic approach to the Kolmogoroff-Smirnov limit theorems. - *Annals of Mathematical Statistics*, 1952, p.277.

- [3] ERDOS - On a theorem of Hsu and Robbins. - *Annals of Mathematical Statistics*. 1949 (20), p.286.
- FORTET - Loi des grands nombres pour des éléments aléatoires généraux. *Proceedings of the international Congress of Mathematicians - Amsterdam (1954)*, Vol.III, p.360.
- [4] FORTET et MOURIER - La convergence de la répartition empirique vers la répartition théorique. - *Annales de L'E.N.S.* (1953), p.267.
- [5] HALMOS - *Measure theory*.
- [8] LOEVE - *Probability theory*.
- [6] PROHOROV - 1 - Probability distributions in functionals spaces. - *USPEHI, Matem. Nauk*. Vol.8, n° 3 (55), 1953, p.165.
- 2 - Convergence de processus aléatoires et théorèmes limites en Calcul des Probabilités. - *Academia NAUK URSS*, 1956.
- [7] REVUZ - *Thèse Paris*.
- SCHWARTZ - Fonction aléatoire du mouvement brownien - *Séminaire Bourbaki* (Février 1958).

TABLE DES MATIÈRES

	Pages
CHAPITRE I - LA LOI DE PROBABILITE ALEATOIRE P_n .	3
A - Convergence simple de P_n^* vers P_x^*	3
B - Convergence uniforme de P_n^* vers P_x^*	8
CHAPITRE II - LA LOI D'ERREUR STATISTIQUE Δ_n	15
A - Limite simple en loi de Δ_n^* dans $C_{P_x}^2$	15
B - Δ^* fonctionnelle linéaire sur D_α^1	17
C - Δ^* fonctionnelle linéaire sur D_α^∞	20
D - Δ^* fonctionnelle linéaire sur D_α^p	21
E - Fonction déterminante de la loi d'erreur.....	23
F - Conclusion : Propriétés de P_n comme estimateur de P_x	26
CHAPITRE III - PRODUITS DE REPARTITIONS EMPIRIQUES.	29
A - Convergence simple de $(P_n \times Q_m)^*$ vers $(P_x \times P_y)^*$..	29
B - Convergence uniforme de $(P_n \times Q_m)^*$ vers $(P_x \times P_y)^*$.	30
C - Loi d'erreur.....	32
D - Conclusion.....	32
CHAPITRE IV - LOIS DES GRANDS NOMBRES POUR DES VARIABLES ALEATOIRES INTER-DEPEN- DANTES.....	35
A - Variables aléatoires enchaînées, type A.....	35
B - Variables aléatoires enchaînées, type B.....	53
C - Extension en cas de plusieurs échantillons.....	56
CHAPITRE V - THEOREME CENTRAL LIMITE POUR DES VARIABLES ALEATOIRES INTER-DEPEN- DANTES PLURI-DIMENSIONNELLES.....	63
Fonction déterminante de la loi d'erreur.....	65

CHAPITRE VI - CLASSES D'INVARIANCE DES FONCTIONS DETERMINANTES D'ERREUR DANS R^k	73
A - Fonctions de Fredholm $\Delta(\lambda)$ et $D(\lambda)$ des noyaux Γ et Γ_1	77
B - Solution de l'équation (E_1)	81
C - Classes de lois de Probabilités sur R^k	82
D - Généralisation.....	84
BIBLIOGRAPHIE.....	87

DEUXIEME THESE

Propositions données par la Faculté

PROGRAMMES LINEAIRES EN ENTIERS

Vu et approuvé

Paris, le 8 Octobre 1961

Le Doyen de la Faculté des Sciences,
J. PERES

Vu

et Permis d'imprimer
Le Recteur de l'Académie
de Paris,
Jean SARRAILH

